

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA**

Lorena Tasca Larré

**Modelagem dos Domínios da Saúde e da Violência na FrameNet Brasil para
identificação de casos de Violência de Gênero**

Juiz de Fora
Agosto 2024

Lorena Tasca Larré

**Modelagem dos Domínios da Saúde e da Violência na FrameNet Brasil para
identificação de casos de Violência de Gênero**

Dissertação de mestrado submetida ao Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora, como requisito parcial à obtenção do título de Mestre em Linguística.

Orientador: Prof. Dr. Tiago Torrent

Juiz de Fora
Agosto 2024

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Tasca Larré, Lorena .

Modelagem dos Domínios da Saúde e da Violência na FrameNet Brasil para identificação de casos de Violência de Gênero / Lorena Tasca Larré. -- 2024.

99 p.

Orientador: Tiago Timponi Torrent

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Letras. Programa de Pós-Graduação em Linguística, 2024.

1. Semântica de Frames. 2. Saúde. 3. Violência. 4. Violência de Gênero. 5. Violência contra a mulher. I. Timponi Torrent, Tiago , orient. II. Título.

Lorena Tasca Larre

Modelagem dos Domínios da Saúde e da Violência na FrameNet Brasil para identificação de casos de Violência de Gênero

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de mestre em linguística. Área de concentração: linguística.

Aprovada em 26 de agosto de 2024.

BANCA EXAMINADORA

Prof(a) Dr(a). Tiago Timponi Torrent - Orientador
Universidade Federal de Juiz de Fora

Prof(a) Dr(a). Carol Martins da Rocha
Universidade Federal de Juiz de Fora

Prof(a) Dr(a). Agelica Terezinha Carmo Rodrigues
Universidade Estadual Paulista

Juiz de Fora, 07/08/2024.



Documento assinado eletronicamente por **Tiago Timponi Torrent, Coordenador(a)**, em 26/08/2024, às 17:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Carol Martins da Rocha, Professor(a)**, em 26/08/2024, às 17:37, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Angelica Terezinha Carmo Rodrigues, Usuário Externo**, em 27/08/2024, às 15:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1907617** e o código CRC **3EAF9A2A**.

Lorena Tasca Larré

**Modelagem dos Domínios da Saúde e da Violência na FrameNet Brasil para
identificação de casos de Violência de Gênero**

Dissertação de mestrado submetida ao Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora, como requisito parcial à obtenção do título de Mestre em Linguística.

Aprovada em 26 de agosto de 2024

BANCA EXAMINADORA

Prof. Dr. Tiago Timponi Torrent
UFJF

Profa. Dra. Carol Martins Rocha
UFJF

Profa. Dra. Angélica Terezinha Carmo Rodrigues
UNESP

AGRADECIMENTOS

Agradeço o apoio dos amigos e familiares, em especial aos amigos que são também colegas pesquisadores e ajudaram no incentivo, na resolução de dúvidas e em um melhor envolvimento com o meio acadêmico. Agradeço à Vital Strategies pela disponibilização dos dados utilizados para a análise, sem os quais não teria sido possível a modelagem de domínios aqui presente e também a toda a equipe da FrameNet Brasil, por todo trabalho em conjunto e todas as trocas de informações indispensáveis à realização de todo o projeto. Agradeço muito ao meu orientador, professor, Dr. Tiago Timponi Torrent, pelo grande incentivo em momentos iniciais de quase desistência, pela paciência e por toda a orientação. Por fim, agradeço à bolsa CAPES processo nº 88887.716453/2022-00 pelo financiamento que tornou este estudo possível.

RESUMO

Dados sobre Violência de Gênero no Brasil e no mundo se mostram alarmantes, de forma que uma em cada três mulheres no mundo já sofreu violência física ou sexual. Pesquisas sugerem que há subnotificação da violência de gênero por diversos fatores, tais como: medo da vítima ou do profissional de saúde, falta de informações sobre a questão e desconhecimento sobre a notificação ou sobre os sistemas de registro. O registro de ocorrências de violência é feito através do SINAN – Sistema de Informação de Agravos de Notificação, no entanto, esse sistema não é integrado a outros sistemas que contêm registros relevantes, como o Sistema e-SUS Atenção Primária a Saúde (e-SUS APS) e o SIM – Sistema de Informação sobre Mortalidade. Este trabalho faz parte de um esforço de linkage de dados entre esses sistemas, no que tange aos campos parametrizados, e de análise semântica baseada em frames, no que se refere aos campos abertos, os quais contêm informações sobre atendimentos médicos e sobre os registros de violência. O linkage de dados permite entender se determinada vítima foi atendida no serviço de saúde antes de uma possível situação de Violência de Gênero. Ademais, os registros do SIM permitem entender se a vítima veio a óbito. A análise semântica baseada em frames procura auxiliar a tarefa de algoritmos computacionais, identificando que tipo de frames foram evocados nos domínios relacionados, que são o da saúde e o da violência. Os dados dos corpora presentes no trabalho são provenientes dos sistemas de saúde (SUS) de São Caetano do Sul – SP; e Recife – PE. O trabalho contribui com a modelagem de 74 frames e 4457 unidades lexicais para os domínios. Os frames são relacionados entre si através de 113 relações tipadas. Um total de 5452 relações qualia conectam unidades lexicais nos domínios. Após a modelagem, a análise de prevalência de frames e unidades lexicais nos corpora do e-SUS APS e do SINAN permite identificar possíveis trajetórias da violência contra a mulher.

Palavras-chave: Semântica de Frames, Violência, Violência de Gênero, Violência contra a mulher, Saúde, SUS, SINAN.

ABSTRACT

Data on gender-based violence in Brazil and around the world are alarming, with one in three women worldwide having suffered physical or sexual violence. Research suggests that gender-based violence is underreported due to several factors, such as fear on the part of the victim or health professional, lack of information on the issue, and lack of knowledge about reporting or registration systems. Incidents of violence are recorded through SINAN – the Notifiable Diseases Information System; however, this system is not integrated with other systems that contain relevant records, such as the e-SUS Primary Health Care System (e-SUS APS) and the SIM – Mortality Information System. This work is part of an effort to link data between these systems, with regard to parameterized fields, and to perform frame-based semantic analysis, with regard to open fields, which contain information on medical care and violence records. Data linkage allows us to understand whether a given victim was treated at the health service before a possible situation of Gender-Based Violence. Furthermore, SIM records allow us to understand whether the victim died. Frame-based semantic analysis seeks to assist the task of computational algorithms by identifying what type of frames were evoked in the related domains, which are health and violence. The data from the corpora present in the work come from the health systems (SUS) of São Caetano do Sul - SP; and Recife - PE. The work contributes to the modeling of 74 frames and 4,457 lexical units for the domains. The frames are related to each other through 113 typed relations. A total of 5,452 qualia relations connect lexical units in the domains. After modeling, the analysis of the prevalence of frames and lexical units in the e-SUS APS and SINAN corpora allows us to identify possible trajectories of violence against women.

Keywords: Frame Semantics, Violence, Gender-based Violence, Violence Against Women, Health, SUS, SINAN.

SUMÁRIO

1 INTRODUÇÃO.....	8
2 ENQUADRAMENTO DA VIOLÊNCIA BASEADA EM GÊNERO PELA SEMÂNTICA DE FRAMES.....	12
2.1 VIOLÊNCIA BASEADA EM GÊNERO.....	12
2.2 SEMÂNTICA DE FRAMES E FRAMENET.....	19
2.2.1 A Semântica de Frames.....	19
2.2.2 A FrameNet.....	23
2.3 SEMÂNTICA DE FRAMES APLICADA AO ESTUDO DA VIOLÊNCIA DE GÊNERO: TRABALHOS ANTERIORES.....	31
3 MATERIAIS E MÉTODOS.....	35
3.1 CORPORA.....	35
3.2 EXTRAÇÃO DE CANDIDATOS A TERMOS.....	39
3.3 CLUSTERIZAÇÃO DOS CANDIDATOS A TERMOS.....	41
3.4 MODELAGEM DOS DOMÍNIOS.....	41
3.5 ANOTAÇÃO DE SENTENÇAS.....	43
4 MODELAGEM DE FRAMES PARA OS DOMÍNIOS DA SAÚDE E DA VIOLÊNCIA	45
4.1 FRAMES DO DOMÍNIO DA SAÚDE.....	45
4.2 FRAMES DO DOMÍNIO DA VIOLÊNCIA.....	64
5 ANÁLISE DOS FENÔMENOS CAPTURADOS PELO MODELO.....	82
5.1 O LÉXICO DA VIOLÊNCIA NOS CAMPOS ABERTOS DOS PRONTUÁRIOS DO E-SUS.....	82
5.1.1 Léxico Evocador dos Frames do Domínio da Saúde.....	83
5.1.2 Léxico Evocador dos Frames do Domínio da Violência.....	85
5.2 O LÉXICO DA VIOLÊNCIA NOS CAMPOS ABERTOS DO SINAN.....	87
6 CONCLUSÃO.....	94
REFERÊNCIAS.....	95
ANEXOS – DICIONÁRIOS DE DADOS DOS SISTEMAS E-SUS APS, SINAN E SIM.....	100

1 INTRODUÇÃO

Esta dissertação apresenta o trabalho de modelagem de frames dos domínios da Saúde e da Violência, o qual faz parte do projeto “Data linkage e análise textual baseada em frames para a identificação de casos candidatos à Violência de Gênero em territórios”. O projeto, uma parceria entre a FrameNet Brasil e a Vital Strategies Brasil, consiste em utilizar análise semântica de campos abertos de sistemas utilizados pelo SUS em hospitais públicos e Unidades Básicas de Saúde (UBS), – como o e-SUS Atenção Primária (e-SUS APS) e o Sistema de Informação de Agravos de Notificação (SINAN) – além de mapeamento de dados de campos parametrizados de tais sistemas de modo a permitir a identificação de territórios propensos à Violência de Gênero.

Os altos índices de Violência de Gênero, bem como as diversas pesquisas (Garbin et al., 2015; Kind et al., 2013) que apontam ainda a subnotificação desses índices, são tratados como assunto de saúde pública, portanto, são relevantes para pesquisas e análises que visem mitigar o problema ou auxiliar em treinamentos profissionais, intervenções e em demais políticas públicas de combate ao problema, além de demonstrar a importância da questão para gestores públicos. Em linhas gerais, o projeto se estrutura conforme a Figura 1.

Dados de prontuários eletrônicos do e-SUS APS, do SINAN e do SIM são coletados nos municípios parceiros do projeto. Tais dados contêm tanto campos parametrizados quanto campos de texto aberto. A equipe da Vital Strategies Brasil, usando metodologia própria, faz o *linkage* dos dados, processo através do qual todos os registros pertencentes a uma mesma pessoa são unificados e ordenados temporalmente. Na sequência, em um esforço conjunto com a FrameNet Brasil, procede-se à anonimização dos dados. A anonimização consiste tanto na remoção dos campos de informações pessoais dos registros, quanto da análise eletrônica dos campos abertos, uma vez que é comum que profissionais de saúde acabem por inserir informações potencialmente identificadoras nos campos de texto. Uma vez que os dados estejam anonimizados, os campos abertos são extraídos e entregues para a equipe da FrameNet Brasil para realização da etapa de análise semântica.

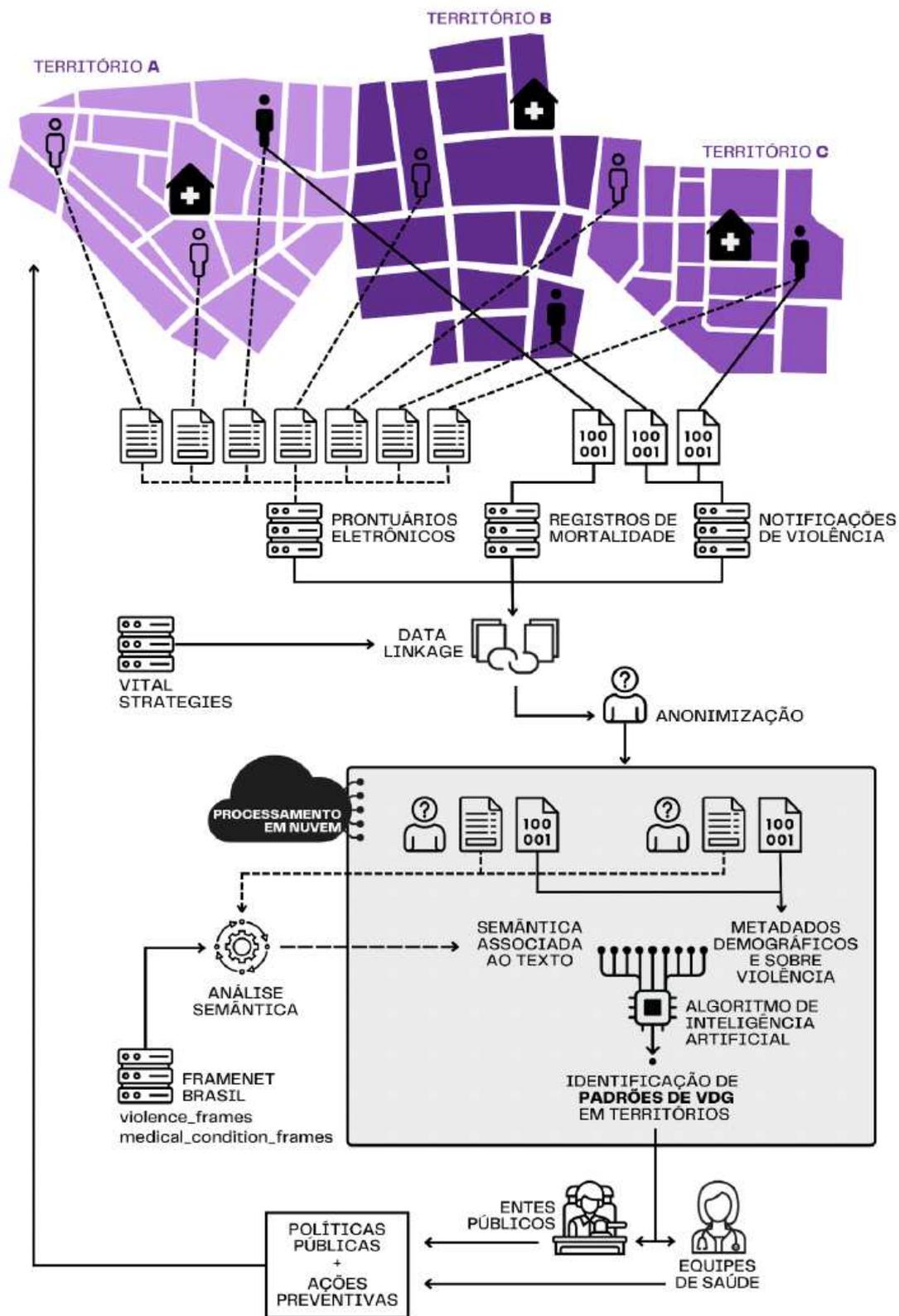


Figura 1: Visão geral do projeto

Fonte: <https://datatostopgbv.org>

Os resultados da análise semântica são então cruzados com os metadados oriundos dos campos parametrizados e um algoritmo de IA usa ambos para identificar padrões de subnotificação de violência contra a mulher. Tais dados, sempre agregados ao nível do território para garantir a segurança das potenciais vítimas, são entregues a gestores públicos e equipes de saúde, que podem, a partir deles, tomar decisões informadas sobre mecanismos de redução de subnotificação e para o enfrentamento da violência.

A hipótese que motiva a análise semântica, baseada em frames, de campos abertos de sistemas de saúde é a de que a vítima que sofre violência doméstica terá procurado os sistemas de saúde e passado por atendimento em algum momento anterior a uma situação mais drástica de violência que venha a ser registrada em um sistema como SINAN, ou ainda, ocasionar um registro de óbito no Sistema de Informações sobre Mortalidade (SIM), cujos dados também são considerados no projeto.

A fração do projeto apresentada neste texto concerne apenas ao tratamento de *corpus* proveniente de campos abertos dos sistemas e-SUS e SINAN, das cidades de São Caetano do Sul/SP e Recife/PE, bem como à posterior análise semântica dos candidatos a termos presentes para a criação de frames, o que compreende a modelagem dos domínios da Saúde e da Violência. Em outro trabalho, Dutra (2024), com dados provenientes do mesmo projeto, aborda uma avaliação de funcionalidade do modelo de análise semântica baseada em frames, apontando, por exemplo, o quanto o modelo auxilia na obtenção de dados mais específicos sobre os casos de Violência de Gênero e na identificação desses casos. Já esta dissertação, que é caracterizada como um trabalho descritivo acerca de dois domínios lexicais do Português Brasileiro, tem por objetivo apresentar os resultados de tal modelagem, bem como uma análise dos frames e unidades lexicais mais recorrentes nos *corpora* analisados.

Ao todo, foram modelados 74 frames e 4457 unidades lexicais. Os frames são relacionados entre si através de 113 relações tipadas. Um total de 5452 relações qualia conectam unidades lexicais nos domínios.

O restante desta dissertação apresenta o trabalho realizado. No capítulo 2, temos o Enquadramento da Violência Baseada em Gênero pela Semântica de Frames, no qual explicamos conceitos relativos à Violência de Gênero, à proposta teórica da Semântica de Frames e à FrameNet. Em Materiais e Métodos, que

compreende o capítulo 3, são relatados as origens e os tratamentos de *corpus*, como comparação com *corpus* genérico e anonimização, esta relativa à proteção de dados sensíveis de usuários do SUS. Ainda no capítulo 3, é apresentada a clusterização de candidatos a termos, os procedimentos padrão utilizados pela FrameNet na modelagem de um domínio e também os relativos à anotação de sentenças para aprendizagem por máquina. No quarto capítulo, demonstramos a modelagem específica dos domínios da saúde e da violência: particularidades do domínio com relação aos frames já existentes e ao *corpus*, a criação de frames próprios do domínio, a modificação de frames já existentes e os resultados em números relativos a este trabalho. Ao final, no capítulo 5, apresentamos as análises dos fenômenos capturados pelo modelo via anotação semântica baseada em frames realizada por máquina, identificando padrões através dos frames e ULs mais frequentes, concluindo no capítulo 6, com as impressões relativas aos resultados obtidos.

2 ENQUADRAMENTO DA VIOLÊNCIA BASEADA EM GÊNERO PELA SEMÂNTICA DE FRAMES

Este capítulo tem por objetivo situar a pesquisa aqui apresentada tanto em termos de seu problema motivador, quanto de seu ancoramento teórico-metodológico. Assim, começamos por discutir o conceito de Violência Baseada em Gênero para, na sequência, apresentar o modelo da Semântica de Frames.

2.1 VIOLÊNCIA BASEADA EM GÊNERO

Primeiramente, faz-se necessário apresentar algumas discordâncias sobre a definição ou a abrangência do termo Violência Baseada em Gênero. Sinalo & Mandolini (2023, p. 2) citam, primeiramente, a definição exposta no website da comissão europeia¹. Essa e outras publicações² em websites de organismos internacionais sobre Violência Baseada em Gênero (ou GBV, do Inglês *Gender-Based Violence*) sugerem a definição do termo como qualquer violência que ocorra contra qualquer pessoa ou determinado grupo de pessoas, baseando-se em seu gênero, o que incluiria violência contra a população LGBTQIA+, por exemplo. Terry (2007), pesquisadora britânica dedicada a questões de gênero, explica que a Declaração sobre a Eliminação da Violência contra as Mulheres de 1993, elaborada com base em movimentos anteriores, como a Declaração para a Eliminação da Discriminação contra as Mulheres, de 1967 e a década das mulheres, que ocorreu de 1976 a 1985 (Azambuja; Nogueira, 2008) define a VAW (Violência contra mulher) como uma subcategoria da GBV, a qual consiste em:

Qualquer ato de violência baseada no gênero que resulte ou possa resultar em dano ou sofrimento físico, sexual ou psicológico às mulheres, incluindo ameaças de tais atos, coerção ou privações arbitrárias de liberdade, quer ocorram na vida pública ou privada. (Terry, 2007, p. 14, tradução nossa)

Assim, enquanto a VAW afetaria especificamente mulheres, a GBV abrangeria, por exemplo, a violência masculina contra homens homossexuais, o envolvimento de mulheres no fenômeno da expropriação de propriedades e os

¹https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-based-violence/what-gender-based-violence_en#relatedlinks

² <https://eige.europa.eu/gender-based-violence/what-is-gender-based-violence>

espancamentos de alunos do sexo masculino por professores do sexo masculino, situações mencionadas na obra de Terry (2007) como uma “performance de masculinidade agressiva”.

Visando, possivelmente, à desambiguação, a OMS intitula sua seção sobre GBV como "Violência contra Mulher"³ e utiliza a definição das Nações Unidas, que entende GBV, conforme já apresentado acima, como sinônimo da definição de VAW. Simister (2012), autor do livro *Gender Based Violence - Causes and Remedies*, apresenta o conceito de GBV como violência perpetrada exclusivamente contra mulheres. Segundo o autor, a utilização do termo GBV implica assumir dois pressupostos: o primeiro é o de que, considerando a violência entre cônjuges, mais mulheres do que homens são vítimas de violência doméstica; e, por último, alguns homens usam a violência para controlar a sua parceira (Bott et al., 2005, p.3 *apud* Simister, 2012, p. 9), sendo que, na maioria das situações, os números de violência contra mulheres são superiores.

Lopes (2005, p. 159 *apud* Azambuja; Nogueira, 2008, p. 104) afirma que um "rosto feminino" dos direitos humanos aparece nas ausências ou nas violações desses direitos, citando que a maioria dos pobres do mundo são mulheres; a maior parte dos analfabetos são mulheres; a maior parte dos crimes sexuais são praticados contra mulheres; as mulheres e jovens são a maior parte das pessoas traficadas e exploradas sexualmente, entre outros pontos.

No âmbito jurídico, no Brasil, Ávila e Mesquita (2020) analisam o conceito de GBV à luz da aplicabilidade da Lei 11.340/06, conhecida como Lei Maria da Penha, cujo objetivo é:

(...) criar mecanismos para coibir a violência doméstica e familiar contra a mulher, nos termos do § 8º do art. 226 da Constituição Federal e da Convenção sobre a Eliminação de Todas as Formas de Discriminação contra as Mulheres e da Convenção Interamericana para Prevenir, Punir e Erradicar a Violência contra a Mulher (Planalto, 2006).

O artigo discorre sobre as diversas decisões relativas à lei e sobre formas de violência familiar que possam se enquadrar como GBV, tais como a violência por parte de irmão contra irmã, por exemplo.

Ávila e Mesquita (2020) demonstram, também, que a lei serve para expressar, em seu próprio texto, que a violência baseada em gênero é praticada em

³ <https://www.who.int/news-room/fact-sheets/detail/violence-against-women>

uma sociedade marcada pelas desigualdades de poder entre homens e mulheres, ocasionando diversas violências contra estas últimas. Ainda segundo os autores, a lei procura alertar operadores do direito para a existência da violência de gênero no que diz respeito às relações sociais, em geral, e familiares, em especial, sendo que a GBV é o motivador político que baseia o programa normativo que está por trás da elaboração da lei de proteção integral à mulher.

Ainda que existam divergências conceituais e que outras questões relativas a gênero, já mencionadas, sejam igualmente importantes, este estudo tratará a GBV como sinônimo de VAW, tomando como vítimas as mulheres⁴ apenas. É importante salientar também que devido à insuficiência de dados e de tempo hábil para tal, não foi possível fazer deste trabalho uma abordagem interseccional.

Nesse contexto, de acordo com estudos e cálculos elaborados por Simister (2012), pelo menos 1 bilhão de mulheres no mundo já sofreu GBV. A estimativa global publicada pela OMS⁵ aponta que uma em cada três mulheres foi submetida à violência física ou sexual por parte de seu parceiro ou violência sexual por parte de um não-parceiro. Uma em cada quatro mulheres jovens (de 15 a 24 anos), que estiveram em um relacionamento, já terá sofrido violência de seus parceiros por volta dos vinte anos. Conforme Tedros Adhanom, diretor geral da OMS, "*a violência contra as mulheres é endêmica em todos os países e culturas, causando danos a milhões de mulheres e suas famílias*". A prevalência da GBV e seus efeitos nas sociedades humanas apresentam-se como fortes justificativas da relevância deste trabalho.

No que concerne às raízes do problema, Saffioti (2015) demonstra em seu livro que a Violência de Gênero está inserida no patriarcado e que tem raízes sociais e políticas, mencionando o papel das desigualdades sociais e da precarização das relações de trabalho no aumento da insatisfação e do consumo de álcool e drogas, os quais, por sua vez contribuem para aumentar as situações de violência doméstica. Conforme dados expostos pela autora, os episódios de violência doméstica acontecem majoritariamente com pessoas do sexo feminino. No Atlas da Violência de 2024⁶, do IPEA, Cerqueira, Bueno e Lima *et al* (2024) apresentam que,

⁴ Entende-se que o conceito de mulher não abrange somente mulheres cis, porém os dados coletados referem-se majoritariamente a esse grupo, devido à falta de marcação dos campos relativos a identidade de gênero por parte dos profissionais.

⁵ <https://brasil.un.org/pt-br/115652-oms-uma-em-cada-3-mulheres-em-todo-o-mundo-sofre-violencia>

⁶ <https://www.ipea.gov.br/atlasviolencia/arquivos/artigos/7868-atlas-violencia-2024-v11.pdf>

na década de 2012 a 2022, ao menos 48.289 mulheres foram assassinadas no Brasil, sendo que, somente em 2022, foram 3.806 vítimas, o que representaria uma taxa de 3,5 casos para cada grupo de 100 mil mulheres. Segundo os autores, há inclusive um aumento de MVCI (Mortes Violentas por Causa Indeterminada) com relação à violência contra a mulher, considerando que, apenas em 2022, 4.172 mortes violentas de mulheres foram classificadas como MVCI. Segundo o Atlas da Violência de 2024, a maioria dos feminicídios (70%) ocorre em domicílio. Ainda no Brasil, conforme divulgação feita pelo Governo Federal⁷, somente no primeiro semestre de 2022, foram registradas mais de 31 mil denúncias de violência doméstica ou familiar contra mulheres. Em território brasileiro, a notificação de um caso de violência é obrigatória. Os profissionais de saúde devem registrá-lo no Sistema de Informação de Agravos de Notificação (SINAN).

O SINAN, conforme descrição disponível em portal do Governo Federal⁸, é alimentado, principalmente, pela notificação e investigação de casos de doenças e agravos que constam da lista nacional de doenças de notificação compulsória, mas é facultado a estados e municípios incluir outros problemas de saúde importantes que constem em sua região. Ainda conforme o site, a utilização efetiva do sistema permite a realização do diagnóstico dinâmico da ocorrência de um evento na população, podendo fornecer subsídios para explicações causais dos agravos de notificação compulsória, além de vir a indicar riscos aos quais as pessoas estão sujeitas, contribuindo, assim, para a identificação da realidade epidemiológica de determinada área geográfica.

Muitos pesquisadores (Garbin et al., 2015; Kind et al., 2013) apontam que há subnotificação de casos de GBV, devido ao fato de muitas mulheres não procurarem os serviços de saúde quando sofrem violência. Além do mais, mesmo quando há procura, há dificuldade por parte dos profissionais de saúde em realizar o registro, devido a vários fatores, tais como sobrecarga, medo de retaliação por parte dos agressores, falta de informações adequadas sobre a importância da denúncia e também à dificuldade de identificar lesões e outros sintomas como indicativos de que a paciente sofreu violência. Questões burocráticas como problemas no diagnóstico e na identificação dos casos, complexidades das doenças ou agravos,

⁷<https://www.gov.br/mdh/pt-br/assuntos/noticias/2022/eleicoes-2022-periodo-eleitoral/brasil-tem-mais-de-31-mil-denuncias-violencia-contra-as-mulheres-no-contexto-de-violencia-domestica-ou-familiar>

⁸ Fonte: <https://portalsinan.saude.gov.br/>

rotinas e protocolos dos serviços, capacidade técnica dos recursos humanos, bem como a falta de conhecimento adequado sobre formas corretas de preenchimento e a frequência correta da alimentação de dados no sistema foram apontados como causas da subnotificação no SINAN.

A notificação no SINAN é feita através de uma Ficha Individual de Notificação (FIN), que é preenchida pelas unidades assistenciais para cada paciente quando da suspeita da ocorrência de problema de saúde de notificação compulsória ou de interesse nacional, estadual ou municipal. A ficha contém sessenta e nove itens para preenchimento, que vão desde dados pessoais completos do paciente a detalhes minuciosos do tipo de agravo, que pode se relacionar a doenças e evidências epidemiológicas ou com as situações de violência, com as quais trabalharemos. O Dicionário de Dados do SINAN pode ser visto no Anexo I.

Para além do SINAN, o sistema público de saúde no Brasil conta, ainda, com outras bases de dados nas quais podem ser inseridas informações decorrentes de episódios de GBV. São elas: o Sistema e-SUS Atenção Primária à Saúde, o Sistema de Informação de Mortalidade (SIM) e os sistemas de prontuários eletrônicos usados nos serviços de saúde em nível municipal e estadual.

O Sistema e-SUS Atenção Primária à Saúde (e-SUS APS) foi desenvolvido a partir da década de noventa, após a Lei nº 8.080, que dispõe sobre a organização do Sistema Único de Saúde (SUS) e o planejamento da saúde nos municípios. Seu objetivo é o cadastro único do paciente, reunindo informações sobre consultas médicas e odontológicas realizadas em domicílio ou no próprio posto de saúde. O sistema ainda coleta informações diversas relativas às condições de vida da população, como esgoto encanado, renda familiar, número de filhos, entre outros. Seu dicionário de dados conta com 55 itens para preenchimento, incluindo campos parametrizados, como nome, idade, raça e sexo e os campos abertos com observações sobre o tipo de atendimento, estes que servirão para a análise linguística e anotação semântica baseada em frames. O campo aberto do qual os dados foram colhidos das cidades de São Caetano do Sul/SP e Recife/PB são, em resumo, as principais queixas dos pacientes que motivaram a ida à UBS. As queixas são, na maioria das vezes, uma única frase que descreve os sintomas do paciente e as condições médicas existentes, bem como há quanto tempo elas ocorrem. Assim como acontece nos sistemas relativos à internação e atendimento, os dados do e-SUS são repassados mensalmente por todos os estabelecimentos de saúde

públicos, conveniados e contratados que realizam atendimentos e consolidados pelos municípios e estados.

Os dados provenientes desse sistema, em sua completude, se fazem essenciais para o *linkage* dos dados com fichas do SINAN, descrito acima, e do SIM, que detalharemos a seguir. A ficha de preenchimento completa do e-SUS está disponível no Anexo II.

Por fim, o SIM (Sistema de Informações sobre Mortalidade) se mostra também importante para nosso projeto, apesar de não termos trabalhado diretamente com análise semântica de seus dados. A finalidade deste sistema é obter dados sobre óbitos no país, fornecendo informações sobre mortalidade para todas as unidades do sistema de saúde. O documento de entrada no sistema é a Declaração de Óbito (DO), que está disponível no Anexo III. A Declaração de Óbito é constituída por nove blocos, com um total de sessenta e duas variáveis, como identificação, endereço, ocorrência e condições e causas do óbito. Os dados do SIM não são escopo desta pesquisa uma vez que são todos parametrizados, sendo usados apenas para definir se um dado óbito foi decorrente ou não de violência.

As fichas de preenchimento que pertencem ao SINAN e ao e-SUS possuem tanto os campos abertos, cujo conteúdo foi extraído para análise semântica, como os campos parametrizados, que trazem informações sobre a identificação do paciente (nome, CPF, identidade, etc), assim como códigos de agravo, datas, e outras informações de identificação, pertinentes ao acompanhamento daquele quadro. É importante notar que a maioria dos campos tem preenchimento parametrizado, ou seja, cabe ao notificador escolher opções a partir de uma lista fechada. Esses campos de preenchimento, que permitem a identificação do paciente, apesar de anonimizados para a análise linguística, auxiliam no posterior *linkage* de dados extraídos, já que não há integração entre os sistemas mencionados. Os fatos de a maioria dos campos de preenchimento presentes serem limitados a códigos relativos ao motivo da internação, por exemplo, ao CID-10, que é o código relativo à classificação das doenças, e de muitos outros campos serem de preenchimento pré-determinado em dígitos, nos direcionamos à análise linguística exclusiva do texto redigido nos campos abertos, que, apesar de terem limitações de caracteres, viabilizam informações adicionais, focadas em cada caso específico, o que possibilita também a análise semântica.

A hipótese que motiva a análise semântica dos dados encontrados nos prontuários médicos conjuntamente com os dados provenientes das notificações de agravo e de mortalidade é a de que a mulher que sofreu violência doméstica, possivelmente, terá procurado atendimento médico, que terá sido registrado em prontuários eletrônicos da atenção básica no e-SUS que, posteriormente, caso tenha infortunadamente vindo a óbito, haverá, por conseguinte, um registro no SIM. Levando em conta que esses sistemas não são integrados pelo governo, o projeto “Data linkage e análise textual baseada em frames para a identificação de casos de Violência de Gênero em territórios” visa utilizar essa integração dos dados obtidos por esses sistemas, bem como análise textual baseada em *frames* para identificar padrões que demonstrem a frequência e a probabilidade de ocorrer violência de gênero em determinado território. Dessa maneira, caso um registro de óbito de uma vítima, no SIM, conste, por exemplo, como “Traumatismo Craniano”, pode haver um outro registro no SINAN, em campo aberto, explicando que a vítima foi “empurrada escada abaixo”. Os registros da atenção básica, mostrarão, possivelmente, que esta paciente já esteve em atendimento médico. A análise semântica baseada em frames, por sua vez, facilita a contagem de termos relacionados a cada domínio, sendo que a relação de sentido entre esses termos pode nos remeter às cenas de violência e aos tipos de atendimentos médicos que as precedem, com sintomas, condições de saúde e diagnósticos diversos. Assim, a estratégia de análise textual desses campos, além da integração desses dados, se faz essencial para o estudo.

Considerando que há protocolo de anonimização dos dados obtidos nos *corpora* e que a identificação dos padrões de GBV é tratada somente no nível dos territórios onde vivem as possíveis vítimas, o principal objetivo do projeto e da dissertação seria contribuir para elaboração de estudos e capacitação de profissionais e governantes, motivando a aplicação de políticas públicas que visem aumentar sensibilização sobre o tema e eventualmente possam reduzir GBV. Os detalhes relativos à anonimização dos dados estarão disponíveis na seção 3.

Nesse sentido, passamos, a seguir, a apresentar o modelo que será utilizado para a análise dos campos abertos: a Semântica de Frames, conforme implementada pela FrameNet.

2.2 SEMÂNTICA DE FRAMES E FRAMENET

A partir da década de 70, pesquisadores que antes eram voltados para o gerativismo, como Lakoff e Fillmore, procuraram demonstrar a importância da semântica e da pragmática para os estudos cognitivos. Passou-se a discordar, então, da proposta gerativista de uma faculdade da linguagem autônoma. Na abordagem proposta pela linguística gerativa (Chomsky, 1965), haveria um módulo da linguagem independente e dissociado dos demais processos cognitivos, sendo que esse sistema seria responsável por associar um *input* externo a formas gramaticais já preexistentes na faculdade de linguagem. O foco de pesquisa seria apenas a *competência*, sem que o *desempenho* interferisse significativamente nos processos relacionados à linguagem. Em discordância com o programa gerativista, que considera a língua como um aparato inato na mente, a Linguística Cognitiva leva em conta a experiência corpórea e os vários processos cognitivos relacionados a essa experiência para a construção da linguagem e da significação.

A corrente cognitivista vem se desenvolvendo como um "arquipélago" de teorias que compartilham entre si princípios fundamentais, como a corporificação, a noção de rede, a perspectivização e a ideia de enquadramento (Geeraerts, 2006). Dentre tais teorias, a Semântica de Frames, fundada por Charles Fillmore, é um dos modelos mais prevalentes, dentre os cognitivistas, para a análise do continuum semântico-pragmático. Ela é também o modelo que embasa as análises realizadas nesta dissertação.

2.2.1 A Semântica de Frames

Alinhada aos pressupostos da Linguística Cognitiva, a Semântica de Frames é um programa de pesquisa em semântica empírica e um modelo descritivo proposto por Charles Fillmore (1982). Conforme o autor, um frame pode ser definido como:

qualquer sistema de conceitos relacionados de tal forma que, para entender qualquer um deles, você tem de entender toda a estrutura na qual eles se enquadram; quando um dos elementos dessa estrutura é introduzido dentro de um texto ou de uma

conversa, todos os outros elementos são automaticamente disponibilizados. (Fillmore, 1982, p.111).⁹

De acordo com Fillmore (1982), o objetivo de uma Semântica de Frames envolve uma perspectiva diferente para enxergar o significado das palavras: as relações entre sentidos e uma análise mais profunda sobre eles e sobre os elementos que os constituem, assim como todos os princípios utilizados na criação de novas palavras por uma comunidade, por exemplo, auxiliam na obtenção de um significado ancorado na experiência humana.

Dessa maneira, a Semântica de Frames está mais relacionada à semântica empírica do que à formal, valorizando mais as continuidades do que as descontinuidades entre linguagem e experiência. Uma das perguntas exemplificadas por Fillmore como relevantes para o programa é: “Quais são as categorias da experiência codificadas pelos membros dessa comunidade de fala por meio das escolhas linguísticas que eles fazem ao falar?” (Fillmore, 1982). Assim, citando o autor:

(...) as palavras representam categorizações de experiências, sendo que cada categoria se baseia em uma situação motivadora que ocorre em um determinado contexto de conhecimento e experiência. A semântica de frames é compreendida como um esforço para entender as razões que levam uma determinada comunidade de fala a criar determinada categoria representada pela palavra, bem como para explicar o significado da palavra demonstrando e esclarecendo essas razões. (Fillmore, 1982, p.112)

A trajetória Fillmoreana que resultará na proposição da Semântica de Frames tem início ainda na sua graduação. As primeiras experiências de Fillmore com a pesquisa envolveram a elaboração de um dicionário de latim. Nesse projeto, sua função era a de construir – à época, manualmente – concordanceadores para verbos latins. Isso fez com que ele voltasse sua atenção para as propriedades semânticas dos verbos relativizadas às valências verbais, ou seja, a relação entre o verbo e os constituintes obrigatórios para a realização da oração (Fillmore, 2012).

Após o contato com outras línguas, como o árabe, o russo e o japonês e diversos trabalhos envolvendo estudos em linguística formal, linguística baseada em *corpus*, ensino de inglês para estrangeiros, entre outros (Fillmore, 2012), o autor

⁹ A tradução dos trechos de Fillmore (1982) citados nesta dissertação foi realizada por Silva (2009).

pôde adquirir uma visão mais elaborada da relação entre os itens lexicais e a organização sintática, que definia seus respectivos comportamentos na língua. Em 1965, com o lançamento do *Aspects of the Theory of Syntax* (Chomsky, 1965), que instaurava um novo padrão analítico na linguística, Fillmore chegou a trabalhar, em Michigan, com estudos em sintaxe. Entretanto, mais adiante, passa a se envolver com pesquisas relativas à gramática de dependência (Tesnière, 1959, *apud* Fillmore, 2012), quando percebeu alguns problemas estruturais relacionados à dependência de constituintes e outras situações, denominadas “*erros embriológicos*” (Tesnière, 1959, *apud* Fillmore, 2012). Estruturas de *gapping*, presentes em sentenças como “*João gosta de maçãs e Maria de laranjas*” chamaram a atenção do linguista (Fillmore, 2012), uma vez que um único núcleo – o verbo *gostar* – teria dois especificadores e dois complementos dos mesmos tipos.

O contexto supracitado se mostrou relevante para a percepção do autor no que concerne principalmente aos predicados e ao significado dos termos utilizados em uma sentença, configurando-se a estrutura argumental como ancorada nas possibilidades sintático-semânticas dos itens lexicais.

Dessa maneira, notou-se a importância dos papéis semânticos dos argumentos dos verbos. A partir da evolução destes estudos, verificou-se também que uma teoria baseada apenas em traços não se mostraria suficiente para descrever todos os papéis semânticos de todos os argumentos de um predicado. Questionando a separação clássica imediata de constituintes como sujeito e predicado, Fillmore (1982) estipulou que os verbos tivessem apenas dois tipos básicos de atributos relevantes para sua distribuição dentro das sentenças: o primeiro seria a descrição da valência da estrutura profunda em termos do que ele chamou de “frames de casos” (Fillmore, 1968); o segundo seria uma descrição em termos de propriedades de seleção.

Os “frames de casos” se constituíam como descrições de palavras predicativas que transmitiam informações como, por exemplo, “este verbo ocorre com três sintagmas nominais, sendo que: um deles executa a ação proposta pelo verbo; outro designa a entidade afetada pela ação, acarretando uma mudança de estado; e o terceiro indicaria um objeto cuja manipulação permite que o executor provoque tal mudança de estado” (Fillmore, 1982, p.114). Esses sintagmas poderiam ser classificados em termos de traços como [Agente, Paciente e Instrumento]. Entretanto, ao classificar diversos tipos semânticos de verbos,

percebeu-se que algumas generalizações semânticas se perdiam, de forma que algumas diferenças cruciais entre sentenças não poderiam ser explicadas simplesmente apontando regras sintáticas e papéis temáticos. Assim, instituiu-se a relativização a esquemas de cenas, sendo necessário outro nível independente de estrutura de papéis para a descrição dos verbos. Segundo Fillmore (1982):

Uma das maneiras possíveis para se formular uma explicação mais completa da semântica lexical é associar algum mecanismo para gerar conjuntos de condições de verdade de uma oração a partir das informações semânticas ligadas individualmente a determinados predicados. Porém, me parecia mais proveitoso acreditar que existem estruturas cognitivas maiores, capazes de fornecer um novo nível de noções de papéis semânticos em termos das quais se poderia caracterizar semanticamente domínios lexicais inteiros. (Fillmore, 1982, p. 115)

A noção de domínio lexical se faz especialmente importante para este trabalho. A primeira tentativa de Fillmore para descrever a estrutura cognitiva que se entende por domínio lexical foi feita a partir de um artigo sobre verbos de julgamento (Fillmore, 1971 *apud* Fillmore, 1982). Para o contexto do julgamento, tomando como exemplo verbos como “condenar” ou “acusar”, foi necessário imaginar uma esquematização de cena que seria essencialmente diferente dos “*frames* de casos”. Nessa concepção não seriam imaginados apenas um agente, ou um paciente, mas papéis realmente desempenhados na cena em questão como: um reclamante – ou vítima –, um juiz, um réu, um defensor e um acusador, que teriam comportamentos típicos esperados para o propósito daquela ação:

(...) importa é termos aqui não apenas um grupo de palavras individuais, mas um domínio de vocabulário cujos elementos, de alguma maneira, pressupõem uma esquematização do julgamento e do comportamento humanos, a qual envolve noções de valor, de responsabilidade, de julgamento, etc., de modo que se pode dizer que ninguém será capaz de realmente entender os significados das palavras de tal domínio se não entender as instituições sociais ou as estruturas da experiência que essas palavras pressupõem. (Fillmore, 1982, p.116)

A partir do trabalho com os “verbos de julgamento”, novos domínios foram analisados, como o de “evento comercial”, por exemplo, que também se mostrou importante para demonstrar como verbos de um grupo extenso e importante do inglês poderiam ser considerados semanticamente relacionados uns aos outros em função dos diferentes modos com que “indexavam” ou “evocavam” a mesma cena

(Fillmore, 1982, p. 116). Assim, para o exemplo relativo a “evento comercial” teríamos elementos centrais, como o “vendedor” ou o “comprador” que evocariam outros itens, como “as mercadorias” ou o “dinheiro”, de forma que o verbo “pagar”, enfocaria tanto as ações do comprador como as do vendedor. Outros verbos como “gastar”, “custar”, “cobrar” seriam entendidos como mais periféricos. Fillmore (1982), então, empregou a palavra *frame* para indicar a maneira estruturada por meio da qual as cenas são apresentadas ou lembradas, sendo este responsável por estruturar os significados das palavras, e as palavras, por sua vez, seriam responsáveis por evocar determinado *frame*. Essas estruturas seriam, então, motivadoras das categorias que os falantes intencionam evocar ao descreverem situações que podem ser independentes da situação discursiva real, do contexto conversacional (Fillmore, 1982). Para o trabalho desenvolvido neste projeto, como já mencionado, foram considerados os domínios da saúde e da violência e os *frames* relacionados a esses domínios.

2.2.2 A FrameNet

Tendo como base a Semântica de Frames e sendo idealizada e desenvolvida também pelo professor Charles J. Fillmore e sua equipe, a FrameNet é um recurso lexicográfico que surgiu em 1997, no International Computer Science Institute (ICSI), em Berkeley, e que originalmente aplicou a teoria da Semântica de Frames (Fillmore, 1982) para a análise das possibilidades lexicográficas de itens lexicais em inglês (Fillmore & Baker, 2009). A metodologia é guiada pela teoria da Semântica de Frames (Fillmore, 1982, 1985), e as análises, sustentadas por evidências em *corpora*. Com o desenvolvimento dos resultados, essa ferramenta passou também a ser útil para profissionais ligados à lexicografia, como alternativa na estruturação de dicionários, e para a Linguística Computacional, no desenvolvimento de tarefas ligadas ao Processamento de Língua Natural (PLN).

As análises lexicais realizadas na Berkeley FrameNet são apoiadas em textos autênticos, advindos do British National *Corpus* (BNC), American National *Corpus* (ANC) e Wall Street Journal (WSJ). Os *corpora* proporcionam mais de 100 milhões de palavras que fornecem subsídio para as análises. A opção por utilizar as ferramentas da Linguística de *Corpus*, submetendo cada generalização a dados reais, faz com que a FrameNet reitere a proposição da Semântica de Frames de

fornecer um programa de pesquisa em semântica empírica para a descrição da língua. O banco de dados da FrameNet de Berkeley encontra-se disponível online no site <https://framenet.icsi.berkeley.edu/>.

A FrameNet Brasil é o projeto responsável pela expansão da FrameNet para o português brasileiro, tendo a sede e seu laboratório situados na Universidade Federal de Juiz de Fora (UFJF), na cidade de Juiz de Fora, em Minas Gerais. O projeto busca modelar frames que retratam a experiência humana em uma base de dados linguísticos, além da implementação de um repositório de construções (Constructicon) para o português do Brasil. Para isso, são aplicadas as teorias da Semântica de Frames já mencionadas (Fillmore, 1975; 1977; 1982; 1985).

Apresentados seus aportes teóricos, faz-se necessário nos aprofundarmos nas questões relativas à implementação do modelo. Um dos conceitos que se faz necessário apresentar é o de Unidade Lexical. Conforme Ruppenhofer et al. (2016):

Uma unidade lexical (UL) é um pareamento de uma palavra com um significado. Normalmente, cada sentido de uma palavra polissêmica pertence a um frame semântico diferente, uma estrutura conceitual semelhante a um script que descreve um tipo específico de situação, objeto ou evento junto com seus participantes e adereços. (RUPPENHOFER et. al., 2016. p. 7, tradução nossa)

Assim, considerando a definição acima, mostramos como exemplo o a UL “*disorder.n*” (*desordem.n* ou *transtorno.n*), nas seguintes sentenças:

- (1) “He called on the authorities to stop public disorder.” (*Ele chamou as autoridades para deter a desordem pública*).
- (2) “Alzheimer's disease is a neurological disorder”. (*O mal de Alzheimer é uma desordem neurológica*).

Na sentença (1), temos “desordem” como sinônimo de “bagunça”. Nesse sentido, o frame evocado é o de *Chaos*¹⁰, representado na Figura 2.

O dicionário lexical da FrameNet de Berkeley demonstra, na definição do frame, que as unidades lexicais que o evocam têm a ver com uma Entidade em Estado de ordem ou desordem (ou transtorno). A Entidade que sofre a desordem e o Estado de desordem descrito são Elementos de Frame Nucleares, nos quais nos

¹⁰ Seguindo a notação tipicamente utilizada pelos trabalhos desenvolvidos pelas framenets ao redor do mundo, os nomes dos frames são apresentados em fonte *Courier*.

aprofundaremos mais adiante. Podemos perceber então que, no inglês, o termo “disorder”, pede como papéis uma Entidade e um Estado, que serão designados, no contexto, pelos elementos da valência da UL *disorder.n*, que evoca o frame de Chaos.

Chaos

[Lexical Unit Index](#)

Definition:

The lexical units in this frame have to do with an **Entity** being in a **State** of order or disorder. Often the **Entity** is unexpressed in this frame and is, therefore, marked as a null instantiation.

Soon **the system** was *in* **total CHAOS**.

Fifa had already fined the Brazilians a minor sum for **crowd DISORDER**.

He had a hammer and banged it against the walls to restore **ORDER** but nobody took any notice of him. **INI**

FEs:

Core:

Entity [Ent]

The object or phenomenon which exists in the **State** of disorder. It is common for the **Entity** to be null instantiated in this frame.

Subsequent bomb alerts caused **travelCHAOS** in the London Underground.

Get the treaty signed before there's time for any more **CHAOS INI**

State [St]

The **State** in which the **Entity** exists. Usually the **State** will be the target-denoting noun.

More **CHAOS** awaits the Soviet Union unless it ditches perestroika for a Thatcherite purge of its economy.

Figura 2: Imagem do frame de Chaos, do Inglês.

Fonte: <https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Chaos>

No entanto, na sentença (2), temos o mesmo lexema, um item polissêmico, sendo utilizado para tratar de uma doença. Para esse caso, outra cena será evocada, relativa ao frame *Medical_conditions* (atualizado para *Condições_de_saúde*, no português), com o qual trabalharemos predominantemente relacionado às ULs obtidas no *corpus* relativo ao domínio da saúde (e-SUS). A Figura 3 apresenta o frame *Medical_conditions*.

Na definição de *Medical_conditions*, temos que os termos para este frame tratam de condições ou doenças que pacientes apresentam; sofrem; estão

sendo tratados para; serão curados de ou que causarão seu óbito. Seus elementos de frame nucleares são a doença em si e o paciente. Podemos perceber que a ambiguidade do termo é facilmente desfeita pelo contexto apresentado pelas próprias sentenças e que os papéis desempenhados pelos diferentes elementos de frame confirmam as cenas, ou os frames, totalmente distintos que o termo *disorder.n* evoca. Assim, ainda que significados de termos polissêmicos possam ter origens etimológicas semelhantes e advir de conceituações parecidas, por uma questão metafórica ou metonímica, por exemplo, os termos consistem em Unidades Lexicais distintas que evocam frames distintos.

Medical_conditions

[Lexical Unit Index](#)

Definition:

Words in this frame name medical conditions or diseases that a patient suffers from, is being treated for, may be cured of, or die of. The condition or disease may be described in a variety of ways, including the part or area of the body (**Body part**) affected by the condition (e.g. liver cancer, cardiovascular disease), the **Cause** of the condition (e.g. bacterial meningitis, viral pneumonia), a prominent **Symptom** of the condition (e.g. asymptomatic stenosis, blue ear disease), the **Patient** or population (originally) affected by the condition (e.g. bovine tuberculosis, juvenile diabetes), or the (proper) **Name** used to identify the condition (e.g. Munchausen Syndrome, Lou Gehrig's Disease). Annotation in this frame is done in respect to the name of the condition or disease.

John died of **pancreatic CANCER**.

FEs:

Core:

Ailment []

FN: Any medical problem.

Patient [Pat]

Semantic Type: Living_thing

This FE is used for the patient, entity, or population (originally) affected by the condition or illness.

juvenile **DIABETES**

Dutch Elm **DISEASE**

Legionnaire's **DISEASE**

Figura 3: Imagem do frame *Medical_conditions* na base da FrameNet de Berkeley.

Fonte: https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Medical_conditions

Sendo importante elucidarmos outros conceitos, consideramos outra UL: *dor.n*. Para compreender adequadamente o significado desse item lexical, temos que considerar uma cena, ou frame, onde um Paciente é afetado por uma Condição. Esses dois participantes são necessários para que o frame seja instanciado e,

portanto, são os Elementos de Frame (EFs) nucleares no frame `Condições_em_saúde`.

EFs são atributos usados como etiquetas para as palavras ou sentenças que estão na construção gramatical com as ULs que o evocam. (Fillmore et al. 2003 *apud* Torrent et al., 2018). São os papéis semânticos das entidades envolvidas em cada *frame* (Fillmore et al. 2003 *apud* Torrent et al., 2018).

O frame `Condições_em_saúde` é apresentado na Figura 4.

Condições_em_saúde [@State] [@Health] [@Lexical] [#215]

Definição

Palavras neste frame dão nome a condições médicas ou doenças que um paciente sofre, está sendo tratado para, pode ser curado de, ou morrer de. A condição ou doença pode ser descrita de várias maneiras, incluindo a parte ou a área do corpo (**Parte_do_corpo**) afetada pela condição (por exemplo, câncer de fígado, doença cardiovascular), a **Causa** da condição (por exemplo, meningite bacteriana, pneumonia), um **Sintoma** importante da condição (por exemplo, estenose assintomática, doença da orelha azul), o **Paciente** ou população (originalmente) afetados pela doença (por exemplo, tuberculose bovina, diabetes juvenil) ou o **Nome** (próprio) usado para identificar a condição (por exemplo, síndrome de

Elementos de Frame Nucleares

FE Core:

Condição_em_saúde Condição que necessita de um serviço médico.

Paciente É usado para o paciente, entidade ou população (originalmente) afetada pela condição ou doença.
semantic_type: @living_thing

Figura 4: Frame `Condições_em_saúde`, na Webtool da FrameNet Brasil.

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Podemos verificar, na imagem, a definição do frame, com os detalhamentos sobre a cena descrita por ele e os Elementos de frame nucleares que viabilizam sua instanciação. Conforme a Figura 4, as palavras neste frame descrevem uma situação na qual um `Paciente` apresenta uma `Condição_em_saúde`. A `Parte_do_corpo`, que normalmente é mais diretamente afetada, pode ser citada e faz parte dos elementos de frame não-nucleares, assim como a `Causa` da dor e outros `Sintomas` que podem estar relacionados. Para o frame exemplificado, há outros exemplos de elementos não-nucleares, como o `Grau` de intensidade, ou o `Nome` da doença.

Quando modelamos frames em uma rede (framenet), podemos estabelecer relações entre os frames recém criados e outros frames já previamente cadastrados. Trata-se das relações Frame-a-Frame (Ruppenhofer et al., 2016). Algumas das relações frame-a-frame (Torrent et al., 2022) podem ser verificadas no *Frame grapher* na Figura 5 e serão detalhadas mais adiante.

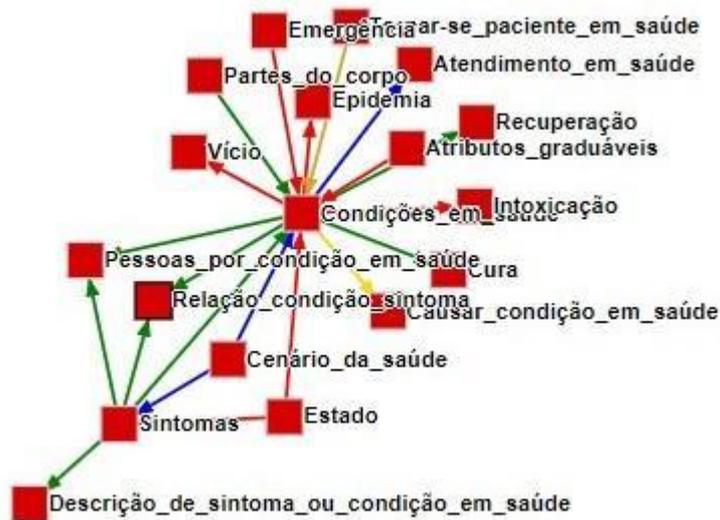


Figura 5: Frame Grapher.

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Em um gráfico de frames e relações (*Frame Grapher*), cada uma dessas relações possui uma cor específica, como podemos verificar na legenda, disponível na Figura 6.

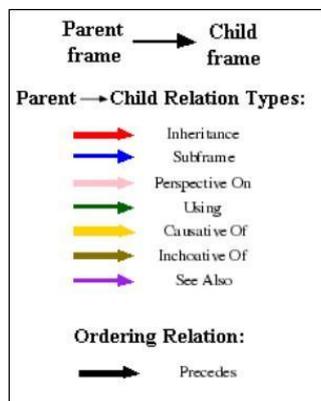


Figura 6: Legenda de relações frame a frame da FrameNet Berkeley.

Fonte: <https://framenet.icsi.berkeley.edu/FrameGrapher>

Para o frame de `Condições_em_saúde`, exemplificado na Figura 4, as cenas de fundo evocadas pelas ULs são conectadas entre si por meio de uma série de relações que modelam:

- a. via *Herança*: que este frame herda e especifica elementos de *frame* de `Estado`, indicando uma situação estável vivenciada por uma Entidade;

- b. via *Subframe*: que as *Condições_em_saúde* consistem em um frame que faz parte do *Cenário_da_saúde* e que frequentemente um *Atendimento_em_saúde*, **subframe das** *Condições_em_saúde*, será requisitado;
- c. via *Uso*: que as *Condições_em_saúde* implicam diretamente em uma *Pessoa_por_condição_de_saúde*, que passará por procedimentos, buscando uma *Recuperação* e posterior *Cura*, sendo então *usado* por esses frames na construção da cena relacionada, bem como *usando* outros frames como *Sintomas* e *Partes_do_corpo*;
- d. via *Causativo_de*: que o frame *Causar_condição_em_saúde*, que representa um processo sistemático evidente de causa-efeito, é a versão causativa que resulta no estado descrito pelo frame *Condições_em_saúde*;
- e. via *Incoativo_de*: que o frame de *Tornar-se_paciente_em_saúde*, caracterizando a situação em que a condição em saúde causa o início do diagnóstico que modifica a pessoa com determinada condição em saúde para paciente, é aquele que indica a entrada do Paciente no estado descrito pelo frame de *Condições_em_saúde*.

As relações de *Perspectiva*, *Precedência* e *Veja* também não são demonstráveis através do frame de *Condições_em_saúde*, mas a relação de *Precedência* está presente em frames como *Atendimento_em_saúde*, que precede o frame *Diagnosticar*, por exemplo. Além das relações frame a frame, outras relações são imprescindíveis para relacionarmos frames, domínios e relativizarmos apropriadamente as unidades lexicais às cenas evocadas. Um exemplo dessas relações, são as relações *Qualia*.

Qualia são entendidos como papéis, estruturas ou relações propostas por Pustejovsky (1995), que descrevem aspectos únicos do significado de uma palavra, ou seja, representam os diversos modos de predicação de um item lexical.

Os *qualia* codificam aspectos do significado de uma palavra que muitas vezes são atribuídos como conhecimento do mundo pelas teorias linguísticas contemporâneas, isto é, o conhecimento que temos sobre objetos no mundo devido à experiência humana. (Pustejovsky & Jezek, 2016, p.4, tradução nossa)

O conceito de qualia surgiu, então, da necessidade de formalização de conhecimento para implementação em uso computacional e da obtenção de uma representação do contexto vivenciado a partir de material linguístico, que não viesse somente do significado dos termos em si, mas daquilo que é expresso linguisticamente, de acordo com o que se conhece do senso comum. Esse conhecimento baseado na experiência é tipicamente descrito através de ontologias. No campo da ciência da informação, ontologias são definidas como “uma explicação explícita e formalizada dos termos de um domínio e das relações entre eles, dada alguma conceituação compartilhada” (Grubber, 1995 *apud* Torrent et al., 2022, p. 7). Apesar das limitações de uma interface entre léxicos e ontologias (Huang et al., 2010 *apud* Torrent et al., 2022), o significado agregado às palavras pode ser, em muitos casos, caracterizado como um conceito ontológico, de forma que algumas relações entre significados podem ser repassadas, até certa medida, para as relações entre conceitos em uma ontologia (Torrent et al., 2022).

No que tange às FrameNets, de acordo com Torrent et al. (2022), “Os frames são de inspiração cognitiva e, portanto, relacionam-se com a ideia de classes em uma ontologia”. No entanto, tentativas de ontologizar a FrameNet de Berkeley revelaram “perdas e distorções do objetivo principal” (Torrent et al., 2022). Após isso, foram analisadas diversas questões relativas à viabilidade de se obter uma ontologia lexical, buscando pontos de concordância entre o léxico e as ontologias, de maneira que as relações entre significados fossem derivadas das relações entre os tipos ontológicos (Torrent et al., 2022).

Na FrameNet Brasil, a implementação desse esforço de ontologização se deu através das Relações Qualia Ternárias. Tais relações associam frames da base da FrameNet Brasil a um dos tipos de qualia originalmente propostos na Teoria do Léxico Gerativo (Pustejovsky, 1995), quais sejam:

1. Formal: descreve a categoria básica do item e fornece as informações que distinguem uma entidade dentro de um conjunto maior, no seu domínio semântico;
2. Constitutivo: expressa uma variedade de relações relativas à constituição interna de uma entidade;
3. Télico: diz respeito à função ou propósito típico de uma entidade, ou seja, para que serve a entidade;
4. Agentivo: diz respeito à origem de uma entidade, o seu criador, ou ao seu surgimento. (Torrent et al., 2022 p. 8)

Para o domínio da saúde, por exemplo, a relação de uso entre o frame de Sintomas e o de Condições_em_saúde, exemplificada na Figura 5, não seria capaz de indicar quais ULs do primeiro se relacionam a quais ULs do segundo. Nesse contexto, um subtipo de quale constitutivo, mediado pelo frame Relação_condição_sintoma, é usado para indicar, por exemplo, que a UL *fadiga.n* no frame de Sintomas se relaciona à UL *covid.n* no de Condições_em_saúde, mas não à UL *Alzheimer.n* do mesmo frame. Assim as qualia ternárias especificam os nexos semânticos na base de dados, aumentando a sua granularidade, de forma que tais nexos sejam agrupados de acordo com ULs relacionadas, mas evocadoras de frames diferentes.

Finalizada a seção que abrange o aporte teórico-implimentacional, iniciaremos a seguir a seção relativa a estudos anteriores envolvendo a análise linguística baseada em frames e a violência de gênero.

2.3 SEMÂNTICA DE FRAMES APLICADA AO ESTUDO DA VIOLÊNCIA DE GÊNERO: TRABALHOS ANTERIORES

De forma semelhante ao trabalho que estamos realizando, questões relativas à GBV foram abordadas por Minnema et al. (2021, 2022a, 2022b). Suas pesquisas propuseram uma análise baseada em frames para verificar como eventos socialmente relevantes, tais como GBV, são narrados no discurso midiático.

A investigação preliminar dos dados conta com um grande compilado de notícias de jornais italianos tratando da temática do feminicídio, além de modelos computacionais devidamente treinados para a tarefa de rotulagem semântica baseada em frames – ou, em inglês, *Frame Semantic Parsing* (FSP). O objetivo dos trabalhos é o de verificar quais tipos de expressões linguísticas são mais utilizadas em notícias relacionadas a GBV e quais frames são mais evocados nessas notícias, de modo a identificar o tipo de enquadramento dado pela mídia ao feminicídio.

Como já demonstramos na seção 2.3, o estudo da Semântica de Frames entende que o significado das palavras está relativizado a cenas, que estão ancoradas na experiência humana. Numa situação de compra e venda, por exemplo, podemos ter a perspectiva do comprador, como em (3); ou a do vendedor,

como em (4); ou ainda, dentro da perspectiva do vendedor, uma maior proeminência dos eletrodomésticos, através do uso da construção passiva, como em (5).

- (3) X compra eletrodomésticos de Y
- (4) Y vende eletrodomésticos para X
- (5) Eletrodomésticos são vendidos por Y para X.

Na sentença (3), a ação executada por X apresenta perspectiva e foco evidente no comprador, enquanto, em (4), o mesmo se verifica para o vendedor. Já em (5), a construção passiva traz os eletrodomésticos em maior evidência, mesmo que (4) e (5) estejam evocando a mesma cena. A perspectivização ora para quem compra, ora para quem está vendendo influencia no tipo de frame evocado e no foco e entendimento por parte do interlocutor, de forma que, levando a perspectivização em conta, ao fazer uma campanha publicitária, por exemplo, provavelmente a loja que vende eletrodomésticos priorizaria o foco na loja e nos eletrodomésticos a serem vendidos, deixando a imagem do comprador em segundo plano, a depender do anúncio.

De maneira semelhante à perspectivização relatada acima, verificando as diversas formas como os jornais abordam e noticiam os fatos, outros estudos (Huttenlocher et al., 1968; Bohner, 2001; Gray and Wegner, 2009; Zhou et al., 2021; Meluzzi et al., 2021 *apud* Minnema et al., 2021) demonstram que as expressões linguísticas utilizadas para descrever um evento impactam na percepção do leitor e no significado social daquilo que é divulgado. Trabalhos anteriores em psicolinguística mostram que, em eventos que envolvem violência em qualquer nível, o *background*, ou seja, o plano de fundo linguístico, e o tipo de expressões utilizadas escondem a responsabilidade dos agentes (perpetradores) e promovem a culpabilização da vítima (Minnema et al., 2021).

Voltando ao tema da GBV, nas descrições de feminicídios obtidas dos jornais utilizados como *corpus* nos estudos de Minnema et al. (2022b), o uso de construções sintáticas com vários tipos de transitividade verbal, desde construções transitivas ativas por um lado, passando por construções passivas e causativas e até construções de nominalização; correspondem a vários graus de responsabilidade atribuídos ao perpetrador (masculino). No exemplo ativo/transitivo (*ele a matou*), ou trazendo um exemplo real do português brasileiro, na sentença (6),

torna-se totalmente explícito o envolvimento de um agente ativo. Já no exemplo passivo (*ela foi morta por ele*), como na sentença (7), o evento é acessado através do paciente, desviando a atenção do agente. Além disso, expressões como “o assassinato (da mulher)” ou mesmo “o evento” (construção nominal) colocam ambos os participantes em segundo plano (Minnema et al., 2022b).

(6) Marido mata mulher no interior de SP¹¹

(7) Mulher é morta pelo ex¹²

Apesar de ser possível encontrarmos exemplos em que há o foco no envolvimento explícito de um agente e, portanto, a perspectiva no perpetrador, as sentenças com foco no paciente, ou seja, na vítima, são mais frequentes. Outros estudos de Minnema et al. (2021) demonstram também que, em geral, os papéis semânticos que se referem ao perpetrador são expressos com muito menos frequência do que aqueles que se referem à vítima. Para o frame *Matar*, por exemplo, o mais evocado por sentenças presentes no *corpus* jornalístico estudado, 60% de todos os casos expressam um papel de vítima, enquanto apenas 33% expressam um papel de assassino (Minnema et al., 2021).

A questão da perspectivização se mostra bastante importante, devido a pesquisas como as do Laboratorio Adolecenza e Instituto IARD (Minnema et al., 2022b), que apontam que jovens acreditam em uma responsabilidade compartilhada entre vítima e perpetrador.

Dada a devida importância aos trabalhos de Minnema et al. (2021, 2022a, 2022b) como precursores de uma abordagem relacionada a GBV baseando-se em frames, é necessário destacar as diferenças entre os estudos acima relatados e o nosso, bem como descrever as contribuições relevantes proporcionadas pelo que está sendo realizado em nossas pesquisas. Os *corpora* estudados nesta dissertação buscam informações retiradas diretamente dos atendimentos de saúde pública e registros de violência, visando relacioná-los para obter via frames e dados parametrizados, o que é e o que pode ser alarmante num atendimento primário,

¹¹ Disponível em:

<https://g1.globo.com/sp/ribeirao-preto-franca/noticia/2024/05/02/homem-diz-que-matou-esposa-por-el-a-se-recusar-a-fazer-almoco-em-santo-antonio-da-alegria-sp.ghtml>

¹² Disponível em:

<https://portalcbn Campinas.com.br/2024/05/mulher-e-morta-pelo-ex-no-dic-v-e-homem-e-presos-em-paulinia/>

quando falamos de GBV. Além disso, conseguimos estender nossa pesquisa não somente a feminicídios, apesar de já serem dados de suma importância, mas também a casos de abuso, espancamento, ou tentativas de feminicídio, nas mais variadas demografias, dados os quais as pesquisas baseadas em frames ou PLN ainda não haviam alcançado.

3 MATERIAIS E MÉTODOS

Este capítulo apresenta os *corpora* e métodos utilizados nesta dissertação.

3.1 CORPORA

A Linguística de *Corpus*, descrita por Sardinha (2004), consiste na coleta e análise de *corpora* ou de conjuntos de informações linguísticas textuais, escolhidas criteriosamente com a finalidade de servir para a pesquisa de uma língua ou variedade linguística, dedicando-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador (Sardinha, 2004, p. 3).

O autor detalha a definição de *corpus*, o histórico relacionado ao surgimento de *corpora* para estudos e análises, que, antigamente, dependiam de anotações e categorizações manuais, sendo de maior costume que esses *corpora* se destinassem mais ao ensino de língua e, posteriormente, fossem sendo utilizados em pesquisas que envolvessem descrição. Aos poucos, com a evolução tecnológica, se tornaram computadorizados, manipuláveis via algoritmos e disponibilizados em diferentes bancos de dados, acessíveis pela internet.

Vale ressaltar aqui a definição de *corpus* encontrada na literatura que, segundo o autor, se adequa melhor à sua essência e seu propósito:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise. (Sanchez, 1995, *apud* Sardinha, 2004, p. 18)

Assim, se mostra evidente a importância de que os dados linguísticos coletados se tratem de uma amostra de língua falada ou escrita naturalmente, sem que sejam propositalmente gerados para a pesquisa em questão, que representem a língua em seu uso cotidiano, devendo ter extensão suficiente para a análise a ser feita.

Os *corpora* utilizados neste trabalho pertencem aos domínios da saúde e da violência e tratam-se de textos digitados em campos abertos, provenientes do

Sistema e-SUS Atenção Primária à Saúde e de registros do SINAN (Sistema de Informação de Agravos de Notificação), obtidos dos municípios de Recife, em Pernambuco, e São Caetano do Sul, em São Paulo. Ambos os sistemas citados são importantes ferramentas digitais, utilizadas por hospitais que realizam atendimentos públicos em todo o Brasil.

A razão para a escolha de informações provenientes desses sistemas, já esboçada anteriormente, é a de que os dados relativos às notificações de agravo, do SINAN, apresentam situações do domínio da violência, incluídas as situações de violência baseada em gênero, e os dados de atendimentos em saúde básica, do e-SUS, apresentam as informações relativas ao domínio da saúde, constando as queixas que levaram o paciente à consulta médica e podendo ou não indicar quais delas possam estar relacionadas também a situações de violência ou violência baseada em gênero. Dessa maneira, a união dos *corpora* pode proporcionar uma análise ampla e uma associação de dados que contribua em estudos, percepções e ferramentas tecnológicas, a fim de criar políticas que possam prevenir ou auxiliar no combate à Violência Baseada em Gênero.

Os dados foram disponibilizados pelas prefeituras dos municípios participantes por intermédio da Vital Strategies, organização global de saúde parceira do Laboratório da FrameNet Brasil, para realização do projeto no qual se insere esta dissertação. Apesar de não ser algo frequente, alguns registros presentes nos *corpora* continham outros tipos de informações, como intervenções médicas anteriores, contatos das vítimas e de seus parentes, que, por sua vez, podem conter informações pessoalmente identificáveis (IPI).

Considerando os direitos dos pacientes e das vítimas à privacidade de seus dados, especialmente quando podem ter sofrido Violência Baseada em Gênero, os passos mais importantes na construção dos *corpora* foram, primeiramente, extrair apenas os campos de texto aberto do banco de dados, para que os lexicógrafos não conseguissem acessar os demais dados do sistema, e, posteriormente, remover quaisquer IPI dos textos. Para este último processo, foi desenvolvido, pela equipe da FrameNet Brasil, um protocolo de anonimização.

O protocolo de anonimização tem o objetivo de identificar as IPIs e substituí-las por etiquetas especiais, indicando que tipo de informação estava naquela frase. Por exemplo, a sentença fictícia de exemplo em (8) seria anonimizada como (9).

- (8) Elvis Presley se feriu no Parque Halfeld às 15:32.
- (9) <NAME> se feriu no <LOCATION> às <TIME>.

A primeira etapa neste protocolo é uma demarcação clara do conceito de IPI no escopo deste estudo. Foram consideradas IPI nomes próprios de pessoas, localidades ou organizações, datas, horas e, de maneira mais geral, números (fossem eles telefones, documentos, endereços...).

O protocolo utilizado neste trabalho usa uma combinação de técnicas de maneira iterativa, que são realizadas automaticamente com uso de algoritmos específicos. Esses algoritmos podem ser classificados em três tipos: o primeiro faz uso de redes neurais treinadas para a tarefa de reconhecimento de entidades nomeadas (REN); já o segundo usa um conjunto de dados manualmente definido e técnicas de busca *fuzzy*¹³; o último utiliza expressões regulares.

O primeiro algoritmo alimenta os textos para cinco modelos REN diferentes e registra onde, nas sentenças, cada um desses modelos identificou nomes de pessoas, organizações e locais, verificando ainda trechos das sentenças onde pelo menos duas redes neurais identificaram uma possível IPI e substitui esses trechos por etiquetas especiais de anonimização. É importante que o algoritmo considere se mais de um modelo identificou aquela IPI, porque esses modelos foram treinados para textos de outros domínios, por exemplo, o jurídico, em diversas línguas e de maneira geral, aderentes à norma culta (Baeza-Yates & Navarro, 1999, Souza et al., 2020, Tedeschi et al., 2021, Cunha e Ramalho, 2022). Em contrapartida, os registros de sistemas de saúde pública formam um gênero textual único, em que essas redes neurais não performam tão bem. O algoritmo compensa essa perda de desempenho considerando vários modelos simultaneamente.

O segundo algoritmo utilizado para identificação de IPI é utilizado exclusivamente para a identificação de locais e organizações. Esse algoritmo faz sucessivas buscas *fuzzy* procurando por nomes conhecidos de um determinado território. Esse processo deve ser parametrizado para cada cidade que produziu os registros dos sistemas de saúde. Como o número de hospitais, bairros, empresas e outros locais é limitado dentro de uma cidade, elabora-se uma lista desses locais,

¹³ https://link.springer.com/chapter/10.1007/3-540-61258-0_1

manualmente ou através de buscas na internet. Essa lista serve de parâmetro para o algoritmo, que tenta identificar nos registros dos sistemas de saúde alguma referência a esses locais. Essa busca pode ser realizada de maneira exata, por exemplo, apenas quando a mesma grafia é utilizada para um local na lista e nos registros de saúde ou através de uma busca *fuzzy*, o que evita que o sistema ignore nomes de locais por conta de erros de digitação ou de grafia.

O último algoritmo utilizado para a substituição automática de IPI faz uso de expressões regulares para remover datas, horas e números. As expressões regulares são representações de padrões flexíveis que podem ser buscados em textos. Por exemplo, a expressão regular $\backslash d\{2\}:\backslash d\{2\}$ pode ser utilizada para encontrar horas no formato hh:mm. O metacaractere $\backslash d$ indica que naquela posição deve haver um dígito, enquanto os números dentro de chaves indicam quantas vezes isso deve ocorrer. Utilizando um conjunto de 8 expressões regulares, todas as ocorrências de datas e horas, assim como números são substituídas pelas suas respectivas etiquetas de anonimização.

A etapa semi-automatizada do protocolo de anonimização consiste no uso de listas de frequência de palavras e de nomes em um banco de dados para identificar *tokens* que sejam nomes em potencial. Esse processo faz uso de uma *wordlist* com frequências de palavras do português. Uma segunda lista é criada a partir de um banco de dados de nomes, também computando as frequências, mas, dessa vez, contando nomes e sobrenomes. Um algoritmo é utilizado para identificar cada etiqueta dos sistemas eletrônicos de saúde que são potenciais nomes, por exemplo, que aparecem na base de dados de nomes e sobrenomes. Para cada um desses possíveis nomes, é computado o seguinte score: $S = plm + fnome - fwordlist$, onde plm é a probabilidade de ocorrência da etiqueta naquele contexto, $fnome$ é a frequência relativa da etiqueta na base de dados de nome e $fwordlist$ é a frequência relativa da etiqueta em um *corpus* geral. Intuitivamente, é esperado que nomes ocorram em contextos específicos, por exemplo, após um pronome de tratamento, e que sejam frequentes no banco de dados de nome, bem como pouco frequentes em um *corpus* geral. Todas as etiquetas com score S consideradas muito frequentes são incluídas em uma lista de possíveis nomes que devem ser verificados manualmente e removidos caso sejam IPI.

Por fim, é feita uma verificação manual dos textos anonimizados. No caso de conjuntos de dados menores, os registros são verificados a fim de garantir que todas

as IPIs tenham sido anonimizadas. Para conjuntos maiores, esse procedimento é conduzido sobre amostras dos dados. Caso alguma IPI seja identificada nessa etapa, as mudanças necessárias são feitas, tanto nos algoritmos, quanto nas expressões regulares ou lista de locais, e o processo é conduzido novamente, até que nenhuma IPI seja encontrada no conjunto anonimizado.

Uma vez finalizado o processo de anonimização, os *corpora* resultantes são compostos de cerca de 206 mil palavras, distribuídas em cerca de 43 mil sentenças, para os dados extraídos do e-SUS APS, e cerca de 172 mil palavras, distribuídas em cerca de 16 mil sentenças, para os dados extraídos do SINAN. Somados, os *corpora* se caracterizam, dentro da proposta de Sardinha (2004), como sendo de tamanho médio, o que é adequado para a pesquisa em questão, visto que busca representar domínios específicos do léxico, em gêneros textuais também específicos, quais sejam os prontuários e as fichas de notificação.

3.2 EXTRAÇÃO DE CANDIDATOS A TERMOS

Após tratados para anonimização dos dados, os *corpora* foram importados para a ferramenta Sketch Engine¹⁴, com a qual foi possível extrair uma WordList com os termos mais frequentes. A WordList é uma ferramenta que gera listas de frequência de termos de diversos tipos, como de substantivos, verbos, adjetivos ou outras partes do discurso.

A partir da WordList, foi feita uma compilação e comparação com *corpus* genérico, ou seja, um *corpus* jornalístico, que aborda diversos domínios do conhecimento e que, portanto, é menos enviesado para a frequência de termos específicos do domínio alvo, para que fossem, então, estabelecidos os candidatos a termos dos domínios específicos do *corpus*.

A extração de termos é um método automático de análise de texto para identificar sintagmas que atendem aos critérios de termos, tendo seu uso na tradução e no gerenciamento de terminologia, mas também na análise de texto, onde é usada para modelagem de tópicos, mineração de dados e recuperação de informações de um texto não estruturado. A extração de termos mais adequada é aquela que fornece uma lista limpa que requer muito poucos ajustes manuais, se

¹⁴ <https://www.sketchengine.eu/>

houver. Muitos métodos tradicionais, que dependem principalmente da frequência no texto de foco, podem apenas extrair termos candidatos que o usuário deve examinar e limpar. A limpeza manual pode ser evitada fazendo uso de critérios linguísticos em combinação com estatísticas. Para que um sintagma se qualifique como um termo, ele deve atender a dois critérios: corresponder à estrutura que um termo pode ter na língua e aparecer com mais frequência no texto em foco do que no texto geral.

Em cada idioma, um termo pode ter um formato diferente. Na maioria das situações, o requisito é que um termo seja um sintagma nominal. Por exemplo, no inglês, um termo pode ser composto de substantivos (N), adjetivos (A) e também preposições, então o sintagma deve corresponder a um desses padrões N+N, N of N, A+N, A+A+N, A+N de N, A+N de A+N etc, enquanto a união “preposição + artigo + adjetivo” provavelmente não será considerada um termo.

Se analisarmos textos de jornais e textos de livros sobre contabilidade, em inglês, por exemplo, é provável que encontremos “*income tax*” e “*best way*” em ambos. Ambos os sintagmas correspondem à estrutura de um termo em inglês (N+N e A+N, respectivamente), no entanto, é provável que as frequências sejam diferentes. Embora a frequência de “*best way*” provavelmente seja semelhante em ambos os textos, a frequência de “*income tax*” provavelmente será muito maior nos textos sobre contabilidade. É assim que o sistema pode diferenciar automaticamente um sintagma frequente de um termo e identificar “*income tax*” como um termo.

Para obter a melhor qualidade possível, o texto em foco deve ser marcado primeiro de acordo com as classes de palavra. Isso garantirá que cada sintagma no texto possa ser comparado com as estruturas de termos permitidas. Sequências de palavras contendo classes indesejáveis podem ser facilmente excluídas com o objetivo de produzir uma lista limpa de termos relevantes.

Outro pré-requisito importante é a lematização. Isso garantirá que as frequências dos termos sejam calculadas corretamente, mesmo que ele seja usado de uma forma diferente, por exemplo, a frequência de “*income tax*” e de “*income taxes*” deve ser calculada em conjunto. Isso é primordial, especialmente para idiomas onde cada verbo, substantivo e outras classes de palavras podem ter várias terminações e variações de forma de palavra.

Ao final, foram extraídos 1000¹⁵ candidatos a termos do domínio da saúde e 1000 candidatos a termos do domínio da violência.

3.3 CLUSTERIZAÇÃO DOS CANDIDATOS A TERMOS

Uma vez extraídos os candidatos a termos de cada *corpus*, estes foram agrupados, primeiramente, de forma mais intuitiva em uma planilha, com listas de lexemas simples e compostos (single e multi-word). Esses lexemas foram tratados para que obtivéssemos os lemas, que são estruturas da base de dados da FrameNet cujo propósito é registrar formas de palavras multiword. Os lemas, por sua vez, foram organizados como candidatos a Unidades Lexicais e essas unidades foram agrupadas em categorias de acordo com o frame que possivelmente evocariam. Desta maneira, algumas Unidades Lexicais foram associadas a frames já existentes na base de dados da FrameNet Brasil e a criação de outros novos frames foi discutida entre os linguistas participantes do projeto no qual essa dissertação se insere. A criação de frames será mais detalhada na próxima seção.

Para entendermos melhor a clusterização inicial, tomaremos por exemplo os termos: médico, doente, dolorido, artrite e joelho. São termos diferentes que, ainda que apresentem particularidades quanto ao sentido e não pertençam todos às mesmas classes de palavras, participam de um mesmo domínio semântico, que é o relacionado à saúde. No entanto, é necessário distinguir o tipo de frame evocado por cada um desses lemas citados. Assim, *médico.n* se enquadra no *frame* de *Profissional_médico*, *paciente.n* se relaciona a *Pessoas_por_condição_médica*, *dolorido.a* evoca o frame de *Sintoma*, *artrite.n* pertence ao frame de *Condições_de_saúde*. Por fim, *joelho.n* foi listado no *frame* de *Partes_do_corpo*.

3.4 MODELAGEM DOS DOMÍNIOS

Os processos para modelagem de um domínio específico dependem primordialmente de um análise intensiva do cenário ao qual este domínio se relaciona e dos *corpora* utilizados.

¹⁵ Para além de 1000 candidatos, uma análise preliminar dos dados mostra que o nível de especificidade cai consideravelmente.

Duas metodologias podem ser seguidas para a criação de frames: *bottom-up* e *top-down*. No método *bottom-up*, o anotador parte dos dados para criar o frame. A primeira tarefa, já mencionada, é selecionar um agrupamento de termos que se relacionem semanticamente.

Em um segundo momento, parte-se para a pesquisa dos itens lexicais selecionados nos *corpora*, tendo em mente um possível frame para investigação. Após esse estudo, selecionam-se algumas sentenças com esses itens lexicais com o objetivo de analisar seu respectivo comportamento no que concerne à valência sintático-semântica. Com isso em mãos, o analista procura regularidades tanto semânticas quanto sintáticas que permitam a estruturação de uma situação específica, essa etapa é a definição do frame. Nela, são determinadas as Unidades Lexicais (ULs), previamente selecionadas pelo analista, são especificadas a nuclearidade dos Elementos de Frame (EFs) bem como a relação entre os EFs. Quando o frame estiver estruturado, é possível relacioná-lo com outros, caso exista entre eles alguma relação, essa é a etapa de assinalar as relações entre frames.

No método *top-down*, as primeiras análises são delineadas pela intuição de falante nativo ou conhecedor da língua, o anotador propõe uma definição para o frame, elencando as possíveis Unidades Lexicais que o evocam, designando a nuclearidade de cada elemento, traçando as possíveis relações entre frames e EFs. Em um segundo momento, verifica-se como foi o seu comportamento nos *corpora* para confirmar se essa definição condiz com os dados. Se necessárias, alterações na definição do frame e dos EFs podem ser realizadas.

De acordo com Gamonal (2013), normalmente, o método *bottom-up*, por partir, primeiramente, dos dados, possibilita uma análise mais ancorada na realidade, sendo mais fiel aos dados que o método *top-down*, que cria o frame para depois investigá-lo no *corpus*. Entretanto, como ambos os métodos são submetidos à apreciação em *corpus*, permitindo que os dados encontrados definam os frames e, conseqüentemente, as considerações acerca de seus elementos, esses métodos não se opõem e, por isso, não devem ser comparados com teor de inferioridade e/ou superioridade.

Para elaborar os frames dos domínios da saúde e da violência foram utilizados ambos os métodos descritos. O modelo construído é apresentado no capítulo 4.

3.5 ANOTAÇÃO DE SENTENÇAS

O objetivo da anotação de sentenças está relacionado às informações semânticas que conseguiremos extrair de determinados *corpora* por meio da anotação. A tarefa de anotação consiste em localizar em cada sentença no *corpus* cada UL e associá-la aos frames já modelados, especificando seus Elementos de Frame.

Conforme Pustejovsky & Stubbs (2012):

(...) existem ainda muitas tarefas que computadores não conseguem realizar, particularmente as que estão no domínio de entender linguagem humana. Métodos estatísticos têm provado ser uma forma efetiva de desenvolver soluções para esses problemas, mas as técnicas do aprendizado de máquina (ML – machine learning) têm trabalhado melhor quando os algoritmos são fornecidos com pontos (ou apontamentos) do que é relevante em um conjunto de dados, ao invés de apenas conjuntos massivos de dados aleatórios. Quando discutimos linguagem natural, esses pontos costumam vir na forma de anotações – metadados (dados sobre dados) que fornecem informação adicional sobre um texto em questão. Entretanto, a fim de ensinar um computador efetivamente, é importante dar a ele os dados corretos e, para isso, ter dados suficientes dos quais ele possa aprender sobre.” (Pustejovsky & Stubbs, 2012, tradução nossa)

Dessa forma, a anotação de sentenças por anotadores humanos se justifica para a criação de metadados que auxiliem algoritmos computacionais no processamento de língua natural (PLN). Após a anotação, os algoritmos computacionais demonstram os resultados elencados como relevantes para análise.

A anotação de sentenças para este trabalho foi realizada na plataforma da FrameNet Brasil, a Web Annotation Tool (Web Tool), seguindo parcialmente a metodologia estabelecida por Ruppenhofer et al. (2016). Uma vez identificadas as Unidades Lexicais, é feita a seleção dos frames previamente criados e a anotação dos Elementos de Frame relacionados com a UL anotada. A anotação é baseada na análise de cada Unidade Lexical presente no texto para cada frame evocado e elementos de frame associados. Na Figura 7 temos um exemplo de fragmento de anotação lexicográfica para a UL *derrame.n*.

[119728] ■ AST_MS_APP		NI	Então, no dia seguinte ela deu um derrame e n
Condições_em_saúde.derrame.n			Então, no dia seguinte ela deu um derrame e n
FE	INC		Pac

Figura 7: Anotação lexicográfica em fragmento de sentença.

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/annotation/main>

Na Figura 7, podemos verificar a anotação para o frame de *Condições_em_saúde*, evocado pela UL *derrame.n*. São anotados os Elementos de Frame Paciente e *Condição_em_saúde*, sendo o segundo marcado como incorporado (INC) ao radical da UL evocadora do frame.

Para o projeto no qual esta dissertação se insere, não foram anotadas as outras camadas tipicamente usadas na FrameNet: Função Gramatical, onde listamos as relações gramaticais que cada elemento possui, e Tipos Sintagmáticos, onde definimos o tipo de sintagma através do qual aquele elemento se manifesta. Tais camadas não são relevantes para um trabalho de extração de informações como o aqui proposto, visto que têm a função primordial de registrar, do ponto de vista lexicográfico, o comportamento morfossintático dos valentes da UL.

A anotação das sentenças para o domínio da saúde proporcionou alguns desafios, como identificação de siglas e abreviaturas comumente utilizadas pelos profissionais de saúde, como, por exemplo, “MMII”, indicando membros inferiores e “USG”, indicando ultrassonografia. Tais siglas foram cadastradas como grafias alternativas (Wordforms) para as ULs. Alguns erros ortográficos e de pontuação, assim como convenções de escrita próprias do gênero, tornam algumas anotações distintas do padrão de dados comumente encontrados em uma FrameNet. O impacto dessas distinções é objeto de análise futura, a ser conduzida após o fim do projeto.

4 MODELAGEM DE FRAMES PARA OS DOMÍNIOS DA SAÚDE E DA VIOLÊNCIA

Este capítulo apresenta os frames modelados para o domínio da saúde e da violência, principais produtos desta dissertação.

4.1 FRAMES DO DOMÍNIO DA SAÚDE

Considerando a metodologia descrita no capítulo 3 para a criação dos frames dos domínios, tendo o *corpus* já passado por procedimento de anonimização e sendo comparado a *corpus* genérico para a obtenção dos termos mais relevantes do domínio, foi feito agrupamento desses termos, que foram reduzidos conforme lema e Unidade Lexical, também de acordo com processos já relatados em outras seções. Para o agrupamento, como já relatado, foi utilizada uma planilha, de acordo com novos frames criados ou realocação de Unidades Lexicais para frames que se mostravam mais adequados.

Com relação ao domínio da saúde, a base de dados da FrameNet já contava com um frame de “Cenário_de_interação_médica” derivado do frame homônimo que consta na base de dados da FrameNet Berkeley. A Figura 8 demonstra o frame original.

Medical_interaction_scenario

[Lexical Unit Index](#)

Definition:

A **Patient** interacts with one or more **Medics** (doctors, nurses, and/or EMTs with a responsibility for the **Patient**'s health) so that the **Medics** can determine the health status of the **Patient** and maintain or restore the **Patient**'s health as necessary.

Usually, the **Patient** has an **Affliction**, a medical problem that motivates or necessitates their coming into the care of a **Medic**. At a minimum, the **Patient** has a single interaction with a **Medic**, but usually the **Patient** has a relationship with the **Medic** or group of medical professionals who either have responsibility for helping them with their specific **Affliction** or for their health in general. Medical interactions normally take place at a **Medical center** (a hospital or doctor's office), but in cases where the **Patient** suffers a sudden catastrophic **Affliction**, the **Patient** may be taken into the care of EMTs, who act as their **Medic** until reaching a medical center.

In some cases, the **Body system** that has a problem is mentioned in place of the **Affliction**.

Her three **cardiac PATIENTS** take up all her time.

Our **doctors** provide the best **HEALTHCARE** on the planet.

FEs:

Core:

- | | |
|-------------------------|--|
| Affliction [aff] | The particular disorder that the Patient is suffering from and the the Medic treats. |
| Medic [Prof] | The medical professional who has responsibility for the care of the Patient . |
| Patient [] | The person who is under the care of a Medic or other form of medical professional. |

Figura 8: Frame Medical_interaction_scenario

Fonte: <https://framenet2.icsi.berkeley.edu/>

Podemos verificar na Figura 8 a definição do frame e os Elementos de Frame que o integram. Este frame, inicialmente, era o único relacionado ao domínio, de forma que, quando o assunto da saúde aparecia em determinado *corpus*, a anotação era feita utilizando apenas ele. Verificou-se a necessidade de alterá-lo para *Cenário_da_saúde*, pois, relacionados a esse domínio, outros profissionais, além do médico, interagem com o paciente: enfermeiros, fisioterapeutas, psicólogos, entre outros. Na Figura 9, vemos o frame do *Cenário_da_Saúde*.

Cenário_da_saúde [@Scenário | @Health | @Scenário | #1138]

Definição

Um **Paciente** interage com um ou mais **Profissionais_em_saúde** (médicos, enfermeiros e / ou paramédicos com responsabilidade pela saúde do **Paciente**) para que os Médicos possam determinar o estado de saúde do **Paciente** e manter ou restaurar a saúde do **Paciente**, conforme necessário. Normalmente, o **Paciente** tem uma **Afiliação**, um problema médico que motiva ou exige que eles cheguem aos cuidados de um Médico. No mínimo, o **Paciente** tem uma interação única com um Médico, mas geralmente o **Paciente** tem uma relação com o Médico ou grupo de profissionais médicos que têm a responsabilidade de ajudá-lo com a **Afiliação** específica ou com a saúde dele em geral. Interações médicas normalmente ocorrem em um **Centro_médico** (um hospital ou consultório médico), mas em casos em que o **Paciente** sofre uma **Afiliação** catastrófica repentina, o **Paciente** pode ser levado aos cuidados de paramédicos, que agem como seu Médico até chegar a um centro médico.

Elementos de Frame Nucleares

FE Core:

- Afiliação**: O distúrbio específico do qual o **Paciente** está sofrendo e o Médico trata.
- Paciente**: A pessoa que está sob os cuidados de um Médico ou outra forma de profissional médico.
- Profissional_em_saúde**: O profissional responsável pelo atendimento do **Paciente**.

Figura 9: Frame do *Cenário_da_saúde*.

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

O processo de modificação do frame existente, atendendo a possível criação de futuros subframes e à realidade do domínio, se caracteriza como um processo mais relacionado ao método *top-down*, na medida em que parte da percepção do pesquisador sobre o domínio sendo modelado.

De outro lado, verificou-se, também, a necessidade de criar outros frames relacionados ao cenário que atendessem ao conjunto de candidatos a termos extraídos do *corpus* do e-SUS e à anotação de sentenças. Assim, o procedimento que apresentamos a seguir para a criação de novos frames se relaciona ao método *bottom-up*.

Normalmente, em uma consulta, um paciente se dirige ou é levado ao hospital ou atendimento médico devido a um sintoma ou condição em saúde. Dessa maneira, foi necessário criar um frame de *Atendimento_em_saúde*, que é evocado por Unidades Lexicais como *consulta.n*. Na Figura 10, temos a imagem do frame.

Atendimento_em_saúde [@Health][#1483]

Definição
Um **Paciente** recebe auxílio de um **Profissional de saúde**.

Elementos de Frame Nucleares

FE Core:

- Atendimento_em_saúde** O atendimento feito pelo Profissional_de_saúde.
- Paciente** A Entidade que precisa de atendimento.
- Profissional_de_saúde** O profissional responsável pelo atendimento.

Figura 10: Frame Atendimento_em_saúde

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Analisando a Figura 10, temos os Elementos de Frame nucleares “Paciente” e “Profissional de saúde”. Tais participantes do domínio também foram modelados como frames: respectivamente, *Pessoa_por_condição_em_saúde* e *Profissionais_em_saúde*. É importante mencionar que, verificando os frames acima e os outros que serão apresentados, é notável que vários dos Elementos de Frame que os integram contêm o item “Paciente”. Dessa maneira, sendo frequentemente citado no *corpus* e não se enquadrando adequadamente no frame de *Pessoas_por_estado_transitório*, frame já existente na base de dados da FrameNet Brasil, concluímos que seria necessário criar o frame de *Pessoas_por_condição_em_saúde*. A imagem do frame com sua descrição e elementos pode ser verificada na Figura 11.

Pessoas_por_condição_em_saúde [@People][@Health][@Lexical][#1469]

Definição
Este frame contém palavras para indivíduos em termos de sua **Condição_em_saúde**. A Pessoa é concebida como independente de outros indivíduos específicos com quem eles têm uma relação e independente de sua participação em qualquer atividade particular.

Elementos de Frame Nucleares

FE Core:

- Pessoa** Pessoa

Elementos de Frame Não-Nucleares

- Condição_em_saúde** Condição médica atrelada a pessoa.
- Descrição** O Descritor é uma condição temporária da Pessoa.

Relações

Figura 11: Frame Pessoas_por_condição_em_saúde

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

O frame *Pessoas_por_condição_em_saúde* herda de um frame mais genérico, de *Pessoas*, existente na base de dados da FrameNet Brasil. O novo

frame, mais apropriado para o domínio, usa diversos frames já citados, como os de Sintomas e o de Condições_em_saúde.

Havia, também, na base da FrameNet, o frame de Pessoas_por_vocação. Entretanto, havendo tantas modalidades diferentes de profissionais em saúde correlatos ao domínio e possíveis Elementos de Frames mais específicos relacionados, como a Aflição a ser tratada e a parte do corpo na qual o profissional é especializado para tratamento, verificou-se, portanto, a necessidade da criação de um frame específico para esse tipo de profissional, herdando do frame de Pessoas_por_vocação. Temos, então, o frame Profissionais_em_saúde na Figura 12.

Profissionais_em_saúde [@People] [@Health] [@Lexical] [#230]

Definição
Este frame inclui palavras que nomelam profissionais da área de saúde e está intimamente relacionado ao frame Especialidades_em_saúde. O EF **Tipo** que caracteriza uma sub-especialidade do **profissional** também pode ser expresso.

Elementos de Frame Nucleares

FE Core:
Profissional Identifica o **profissional** médico.

Elementos de Frame Não-Nucleares

Aflição O distúrbio específico que o **Profissional** médico trata. A maioria das palavras no frame indica inerentemente essa informação.

Base_contratual A **Base_contratual** descreve as condições de emprego em relação à permanência, horas por período de tempo ou base de pagamento.

Compensação O pagamento que o **profissional** recebe por trabalhar.

Empregador O **empregador** dá compensação ao **profissional** por trabalhar.

Etnia É o grupo religioso, racial, nacional, sócio-econômico ou cultural ao qual o **Profissional** pertence.

Idade O período de tempo em que o **Profissional** está vivo.

Lugar_de_trabalho O lugar onde o **Profissional** trabalha.

Origem A **Origem** é o lugar onde o **profissional** nasceu ou viveu uma parte importante de sua vida.

Posição É a posição em uma hierarquia que o **Profissional** ocupa dentro de uma organização.

Figura 12: Frame Profissionais_em_saúde

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Um profissional em saúde normalmente possuirá uma especialidade em saúde e será responsável também por uma intervenção em saúde. Considerando isso, dois frames, um para especialidades em saúde e outro para intervenção em saúde foram criados. Eles evocam ULs como *ortopedista.n* e *biópsia.n*, respectivamente. Nas Figuras 13 e 14, temos os respectivos frames.

Especialidades_em_saúde

[@Entity] [@Health] [@Lexical] [#2]

Definição	
Este frame inclui palavras que nomeiam especialidades em saúde e está intimamente relacionado ao frame de Profissionais_em_saúde. O EF Tipo que caracteriza uma subárea em uma Especialidade também pode ser expresso.	
Elementos de Frame Nucleares	
FE Core:	
Especialidade	Identifica a Especialidade médica que alguém pratica.
Elementos de Frame Não-Nucleares	
Afiliação	A doença ou distúrbio que um especialista tenta tratar. Este conceito é normalmente incorporado ao nome da Especialidade .
Sistema_do_corpo	A sub-parte ou subsistema do corpo que é estudado. Isso normalmente é incorporado nas ULs do frame.
Tipo	Identifica o Tipo de Especialidade médica.

Figura 13: Frame Especialidades_em_saúde

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Intervenção_em_saúde

[@Action] [@Health] [@Lexical] [#1155]

Definição	
As intervenções baseadas em procedimentos ou em medicina são usadas no Paciente para tentar ativar uma Condição_médica . Essas intervenções podem ter uma Frequência_de_sucesso e Efeitos_colaterais . Esse frame difere-se de Cura, já que este frame lida apenas com tentativas de aliviar uma Condição_médica , enquanto Cura lida com situações em que a afiliação ou a Condição_médica foi curada.	
Elementos de Frame Nucleares	
FE Core:	
Condição_em_saúde	Uma descrição holística do estado médico do paciente (ou uma parte do estado do paciente), que pode ou não indicar a causa de um desvio do normal.
Intervenção	Um medicamento ou procedimento é administrado ou realizado para tratar uma Condição_médica .
Profissional_médico	O indivíduo ou equipe que tenta melhorar a condição médica do Paciente .
Resultado	A consequência da Intervenção .
Elementos de Frame Não-Nucleares	
Efeitos_colaterais	Efeitos da Intervenção que são não-intencionais ou indesejáveis.
Extensão	O grau em que um Intervenção afetou a Condição_médica ou os sintomas.
Frequência_de_sucesso	A regularidade com a qual uma Intervenção é eficaz para tratar uma Condição_médica ou os sintomas.
Paciente	Indivíduos que recebem Intervenção médica por causa de sua Condição_médica .

Figura 14: Frame Intervenção_em_saúde

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Como exemplo de intervenção em saúde, temos vários procedimentos, como solicitação de exames (que será descrita mais adiante), receituário, encaminhamento para profissional especializado, etc. Um outro exemplo de intervenção, comum, mas que possui particularidades por se tratar de um objeto e, portanto, herdar do frame *Artefato*, é a *Prótese*. Apesar de, no momento, ser evocada apenas pela UL homônima, acreditou-se necessário elaborar frame específico, devido a particularidade do item e sua função essencial no domínio. Segue Figura 15 representando o frame.

Prótese

Definição
Este frame faz referência à dispositivos ou artefatos que substituem alguma parte do corpo humano.
Elementos de Frame Nucleares
<p>FE Core:</p> <p>Parte_do_corpo Parte do corpo humano que é substituída por uma prótese.</p> <p>Prótese Estrutura artificial que tem como função substituir uma Parte_do_corpo.</p>

Figura 15: Frame *Prótese*

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Percebendo-se a quantidade de ocorrências no *corpus* como “Paciente refere X [Condição_em_Saúde]” ou “Paciente refere W [Sintoma] e “Y [Sintoma] em Z [Parte_do_corpo]”, foi adaptado o frame *Health_Conditions*, proveniente da *FrameNet Berkeley* para o frame de *Condições_em_saúde*, respeitando o raciocínio anterior, de que se trata de um domínio da saúde e não necessariamente um domínio médico. Assim, levando em conta os exemplos acima, onde as sentenças diversas descrevem vários tipos de sintomas, um frame de *Sintomas* foi criado. Vale ressaltar a diferença entre o frame de *Condições_em_saúde*, mencionado acima e o frame de *Sintomas* criado para o domínio a partir dos seguintes exemplos de Unidades Lexicais: a *coriza.n*, evocará o frame de *Sintomas*, enquanto *rinite.n*, evoca o frame de *Condições_em_saúde*. O frame *Condições_em_saúde* já foi demonstrado anteriormente na seção 2.2.2, Figura 4. Na Figura 16, a seguir, temos a imagem do frame *Sintomas*, sua definição e elementos de frame.

Sintomas [@State] [@Health] [@Lexical] [#1464]

Definição
Palavras neste frame dão nome a Sintomas apresentados por um Paciente.

Elementos de Frame Nucleares

FE Core:

Paciente É usado para a Entidade afetada.

Sintoma Alteração de ordem física ou psicológica experienciada por uma Entidade.

Elementos de Frame Não-Nucleares

Condição_em_saúde Condição relacionada ao sintoma.

Descrição Identifica uma característica ou descrição do sintoma.

Duração É usado para expressar a duração do sintoma.

Frequência A frequência com que o sintoma é experienciado.

Intensidade É usado para descrever a intensidade do sintoma.

Parte_do_corpo Parte ou área do corpo afetada pelo sintoma.

Figura 16: Frame Sintomas

Fonte: <https://webtool.framenetbr.ufff.br/index.php/webtool/main#>

Naturalmente, um sintoma ou uma condição em saúde, que necessitam de tratamento ou intervenção, são causados por fatores diversos. Esses fatores podem incluir patógenos e/ou alérgenos ou outras situações, de tal forma que foi necessário um frame que abrangesse as várias formas de acometer a saúde humana ou adquirir alguma condição especial. Isso posto, foi criado o frame de `Causar_condição_em_saúde`, que está relacionado a ULs como `ingerir.v` ou `acometimento.n`. Na Figura 17, temos o frame correspondente.

Causar_condição_em_saúde

Definição
A ação de causar uma **Condição_em_saúde**, intencionalmente ou não.

Elementos de Frame Nucleares

FE Core:

Causa_da_condição_em_saúde A causa da **Condição_em_saúde** que afeta o Paciente.

Condição_em_saúde Condição que necessita de um serviço médico.

Paciente É usado para a Entidade afetada.

Figura 17: Frame Causar_condicao_em_Saúde

Fonte: <https://webtool.framenetbr.ufff.br/index.php/webtool/main#>

Doenças e sintomas são causados também por diversos tipos de patógenos e alérgenos de maneira que frames homônimos também foram criados para itens relacionados que necessitam de anotação. Então, ULs como *bactéria.n* e *camarão.n* estão relacionadas aos frames de Patógenos e Alérgenos, respectivamente, frames que estão representados nas Figuras 18 e 19.

Patógenos

Definição	
Este frame faz referência a entidades biológicas que podem causar uma doença.	
Elementos de Frame Nucleares	
FE Core:	
Doença	Doença causada pelo Patógeno.
Patógeno	Entidade responsável por causar uma doença.
Elementos de Frame Não-Nucleares	
Descritor	Descrição do Patógeno ou de alguma característica relacionada a ele.
Origem	Origem do Patógeno.

Figura 18: Frame Patógenos

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Alérgenos

Definição	
Este frame diz respeito à substâncias de origem natural que podem causar reações alérgicas.	
Elementos de Frame Nucleares	
FE Core:	
Alérgeno	Substância que pode causar uma reação alérgica.
Elementos de Frame Não-Nucleares	
Descritor	Uma caracterização do alérgeno.
Origem	Origem do alérgeno.

Figura 19: Frame Alérgenos

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Assim, relacionado às situações que causam um determinado sintoma ou condição em saúde, temos o ato de infectar. Com isso, foi trazido para a FrameNet Brasil o frame próprio da FrameNet Berkeley, que é usado pelo frame `Patógenos` e herda de `Causar_condição_em_saúde`: o frame `Infectar` é evocado então por ULs como *infectar.v* ou *infecção.n* e pode ser visto na Figura 20.

Definição	
A ação de espalhar uma Infecção para uma Entidade_Infectada , intencionalmente ou não. A patologia indica que a Causa_da_infecção pode ser pelo ar, carregada pela pele ou cabelo, ou transmitida por várias outras formas de contato.	
Elementos de Frame Nucleares	
FE Core:	
Causa_Infecção	Este EF etiqueta a causa da Infecção que afeta a Entidade_Infectada .
Entidade_Infectada	Este EF etiqueta o recipiente da Infecção .
Infecção	Uma invasão, tipicamente indetectada, de uma entidade ou sistema. A Infecção muitas vezes compromete a estabilidade da entidade. Infecções são patologias comuns para doenças.

Figura 20: Frame Infectar

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Relacionado a sintomas, frame visto anteriormente, já havia um frame na base de dados da FrameNet Berkeley que dava conta de associações entre uma determinada condição em saúde a um sintoma específico. Esse frame foi adaptado também para a FrameNet Brasil como `Relação_condição_sintoma`. Ele é

evocado por ULs como *devido.prep* ou *decorrer.v* que indicam que determinado sintoma, como *dor abdominal.n*, por exemplo, pode decorrer de *apendicite.n*.

É importante notar também a existência constante de relatos que envolvem determinada aflição em uma parte do corpo, então um frame de `Partes_do_Corpo`, que é evocado por Unidades Lexicais como *garganta.n* já existia na base de dados da FrameNet e foi constantemente utilizado. O Frame `Partes_do_corpo` é utilizado por outro frame criado com base em itens lexicais presentes no domínio, que é o de `Substâncias_corporais`. Nas Figuras 21 e 22, apresentam-se os dois frames.



Figura 21: Frame `Partes_do_corpo`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

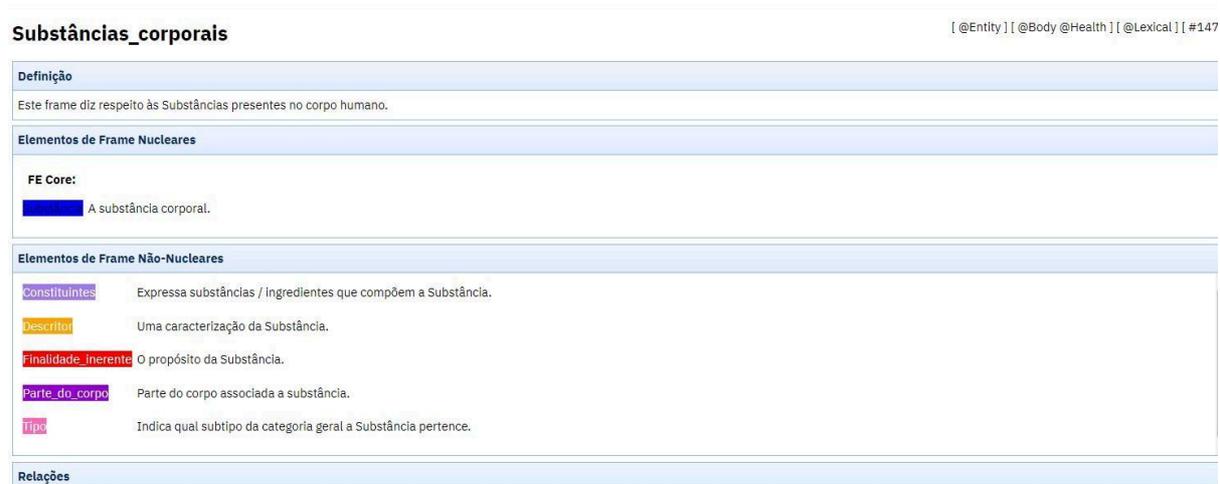


Figura 22: Frame `Substâncias_corporais`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

De forma semelhante ao frame de *Pessoas_por_vocação*, citado mais acima, entre outros frames que foram criados com base em frames preexistentes e adaptados ao domínio, já havia um frame de *Substâncias* na base de dados da FrameNet Brasil que, de forma mais genérica, era evocado por ULs que designavam diversos tipos de fluidos ou outros elementos químicos. No entanto, para este domínio, considerando a expressiva ocorrência de ULs como *sangue.n*, *muco.n*, *bile.n*, considerou-se conveniente criar um frame que abrangesse termos mais relacionados ao tema da saúde.

Ainda relacionado a substâncias ou à frequência de termos no *corpus*, temos muitas vezes menção a medicamentos ingeridos pelos pacientes ou receitados pelos médicos. Nesse aspecto, foi criado o frame de *Medicamentos*, que é evocado por ULs como *analgésico.n* ou *anticoagulante.n*. O frame pode ser visto na Figura 23.

Medicamentos

Definição	
Este frame diz respeito a medicamentos. O Medicamento é geralmente incorporado.	
Elementos de Frame Nucleares	
FE Core:	
Medicamento	Substância utilizada com objetivo de curar ou aliviar um sintoma ou uma condição médica.
Elementos de Frame Não-Nucleares	
Descritor	Descrição do medicamento.
Dosagem	Especifica a dosagem indicada para uso ou aplicação do medicamento.
Modo de administração	Especifica a forma como o medicamento é ou deve ser administrado.
Propósito	Usado para identificar o propósito do medicamento.
Tipo	Usado para especificar o tipo do medicamento.

Figura 23: Frame *Medicamentos*

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Na mesma linha dos medicamentos, eram perceptíveis as diversas menções a materiais utilizados no atendimento em saúde. Portanto, foi criado também um frame de `Materiais_de_uso_em_saúde`. Ele herda do frame `Artefato`, já existente na FrameNet e é evocado por ULs como *agulha.n* e *ampola.n*, respeitando também a especificidade do domínio. O frame de `Materiais_de_uso_em_saúde` pode ser visto na Figura 24.

Materiais_de_uso_em_saúde

Definição	
Este frame inclui palavras para materiais utilizados na área da saúde - por exemplo <code>Instrumentos</code> criados para uma função institucional ou <code>Finalidade</code> .	
Elementos de Frame Nucleares	
FE Core:	
<code>Instrumento</code>	Identifica o <code>Instrumento</code> médico.
semantic_type: @physical_entity	
Elementos de Frame Não-Nucleares	
<code>Finalidade</code>	Ele identifica a <code>Finalidade</code> para a qual um <code>Instrumento</code> é usado.
semantic_type: @state_of_affairs	

Figura 24: Frame `Materiais_de_uso_em_saúde`

Fonte: <https://webtool.framenetbr.ufff.br/index.php/webtool/main#>

Relativamente ao atendimento em saúde e de suma importância, apesar de não ser frequentemente citado nesse *corpus*, encontramos a situação de monitoramento da condição de um paciente. Esse monitoramento é feito através de sinais vitais. Então, um frame de mesmo nome foi criado para dar conta da situação descrita. O frame `Sinais_Vitais` é evocado por ULs como *frequência cardíaca.n* e pode ser visto na Figura 25.

Sinais_vitais

[@Physical_attri

Definição	
Este frame diz respeito aos Sinais Vitais que indicam o estado de saúde de uma Entidade. Eles podem estar relacionados às funções circulatórias, respiratórias, neurais e endócrinas.	
Elementos de Frame Nucleares	
FE Core:	
<code>Paciente</code>	Paciente
<code>Sinal_vital</code>	Indicador do estado de saúde de uma Entidade.
Elementos de Frame Não-Nucleares	
<code>Descricao</code>	Descrição do Sinal_vital.
<code>Grau</code>	Descreve o Grau do Sinal_vital.

Figura 25: Frame `Sinais_Vitais`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Outro frame pouco citado neste *corpus* específico, mas que se mostra importante para o domínio da saúde e que pode vir a ser trabalhado em outras pesquisas, é o frame `Epidemia`. Ele herda de `Condições_em_saúde` e pode abranger várias doenças e situações de saúde pelo mundo. O frame é evocado por ULs como *dengue.n* e pode ser visto na Figura 26.

Epidemia

Definição	
Este frame faz referência ao surgimento de doenças que se espalham com rapidez.	
Elementos de Frame Nucleares	
FE Core:	
Doença	Doença que apresenta um grande número de casos em um curto período de tempo.
Entidade_afetada	Indivíduo(s) afetado(s) pela doença que causou a epidemia.

Figura 26: Frame `Epidemia`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Em muitas situações, há também sentenças com descrições relativas aos sintomas ou às condições em saúde apresentadas, ou ainda, relativas a algum aspecto de alguma parte do corpo. Considerando essas situações, foram criados os frames de `Descrição_de_sintoma_ou_condição_em_saúde` e `Descrição_orientada_pela_parte_do_corpo`, que abarcam ULs como *benigno.a*, para o primeiro frame e *abdominal.a*, para o segundo. Apresentam-se esses frames nas Figuras 27 e 28, respectivamente.

Descrição_de_sintoma_ou_condição_em_saúde

Definição	
Este frame está relacionado a palavras que descrevem Sintomas ou Condições médicas. A Descrição é geralmente incorporada.	
Elementos de Frame Nucleares	
FE Core:	
Condição_em_saúde	Condição que necessita de um serviço médico.
Descrição	Descrição do Sintoma ou da Condição_médica.
Sintoma	Alteração de ordem física ou psicológica experienciada por uma Entidade.
Elementos de Frame Não-Nucleares	
Grau	Identifica o grau ao qual a descrição se aplica.
Parte_do_corpo	A Parte_do_corpo em que a Condição_em_saúde ou o Sintoma se manifesta.

Figura 27: Frame Descrição_de_sintoma_ou_condição_em_saúde

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Descrição_orientada_pela_parte_do_corpo	
Definição	
Este frame está relacionado a palavras que descrevem um estado a partir de uma Parte_do_corpo. Exemplo: Ele está com dor abdominal.	
Elementos de Frame Nucleares	
FE Core:	
Parte_do_corpo	É Parte do corpo.
Elementos de Frame Não-Nucleares	
Descrição	Usado para uma característica ou descrição da Parte_do_corpo.
Subparte_do_corpo	É a Subparte do corpo.

Figura 28: Frame Descrição_orientada_pela_parte_do_corpo

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Ainda relativo à elaboração de frames, situações que demonstravam o relato de exames e hipóteses diagnósticas levaram à criação dos frames Fazer_exame_em_saúde e Diagnosticar que são evocados por ULs como *hemograma.n* e *hipótese diagnóstica.n*, nessa ordem. Esses frames estão descritos nas Figuras 29 e 30.

Fazer_exame_em_saúde

Definição	
Este frame diz respeito ao ato de realizar um Exame_médico.	
Elementos de Frame Nucleares	
FE Core:	
Exame_médico	O exame feito por um Paciente.
Paciente	A Entidade que realiza o Exame_médico.
Elementos de Frame Não-Nucleares	
Descritor	Descrição ou informação adicional sobre o Exame_médico.
Finalidade	O propósito do exame.
Lugar	O Lugar em que o exame é realizado.
Maneira	Qualquer descrição sobre a realização do exame que não estiver coberta por EFs mais específicos.
Profissional_de_saúde	O Profissional_de_saúde responsável pelo pedido ou realização do exame.
Tempo	O momento em que o exame é realizado.

Figura 29: Frame Fazer_exame_em_saúde

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main#>

Diagnosticar

Definição	
Um Profissional_de_saúde apresenta um Diagnóstico para um Paciente.	
Elementos de Frame Nucleares	
FE Core:	
Diagnóstico	Diagnóstico apresentado pelo Profissional_de_saúde.
Paciente	A pessoa que recebe o diagnóstico.
Profissional_de_saúde	O responsável por diagnosticar o Paciente.

Figura 30: Frame Diagnosticar

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main#>

Considerando as diferenças entre diagnosticar e fazer exame bem como a diferença entre fazer exame e entregar um exame já pronto e documentado para

análise do profissional, foi criado também um frame de `Exame_em_saúde`, que está descrito na Figura 31.

Exame_em_saúde

Definição	
Este frame diz respeito à exames médicos solicitados por um <code>Profissional_da_saúde</code> .	
Elementos de Frame Nucleares	
FE Core:	
Exame	A forma física do exame. Pode se tratar do pedido ou do resultado.
Elementos de Frame Não-Nucleares	
Descritor	Descrição ou informação adicional sobre o Exame

Figura 31: Frame `Exame_em_saúde`

Fonte: <https://webtool.framenetbr.ufff.br/index.php/webtool/main#>

Por fim, para toda condição em saúde, envolvendo sintomas, partes do corpo, medicamentos e outras intervenções, poderemos ter uma `Recuperação` ou uma `Cura`. Apesar das diferenças sutis entre os dois eventos, é importante pensarmos que um quadro de recuperação não necessariamente indica a cura completa do paciente, mas claro, já é um passo importante. Disto isto, na base da FrameNet Brasil já havia um frame, originado de Berkeley que era relativo apenas à `Recuperação`, evocando ULs como *cicatrização.n* e *medicado.a*. Tal frame é mostrado na Figura 32.

Recuperação

[@Change] [@Health] [@Lexical]

Definição	
Estas palavras descrevem a recuperação ou cura de um Paciente de uma Aflicção sem referência à Influência de qualquer Tratamento ou Curador (veja o Frame Cura).	
Elementos de Frame Nucleares	
FE Core:	
Aflicção requires: Paciente	Esse EF é o ferimento, dor ou doença vivenciada pelo paciente.
Paciente requires: Aflicção	O Paciente pode ser expresso como o Argumento Externo ou, como em Cura, pode ocorrer como o possuidor da Aflicção (um exemplo de confiação de EFs).
Parte_do_corpo semantic_type: @body_part	Onde a Parte_do_corpo é expressa como um constituinte separado, ela ocorre como um Argumento Externo. A Parte do Corpo é tratada como um Elemento de Frame separado, embora semantic_type : @body_part possa ser vista como metonímica para a Aflicção.
FE Core set(s): {Paciente,Aflicção,Parte_do_corpo}	

Figura 32: Frame Recuperação

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

O frame de Recuperação é causativo do frame de Cura (via relação causativo_de) e o precede (via relação de Precedência). O frame de Cura é evocado por ULs como *alta hospitalar.n* e *resolução.n*.

Cura

[@Action] [@Health] [@Lexical] [#

Definição	
Este frame trata de um Curador tratando e curando uma Aflicção (as lesões, doenças ou dor) do Paciente , algumas vezes mencionando também o uso de um Tratamento ou Medicação , em particular. Esse frame difere de Intervenção_médica, pois este frame lida apenas com casos em que o Paciente é curado da Aflicção , não apenas tratado para a Aflicção .	
Elementos de Frame Nucleares	
FE Core:	
Aflicção	Geralmente é o Objeto SP de um verbo, freqüentemente incorporando o Paciente como um possuidor.
Curador semantic_type: @sentient	É qualquer um que trata ou cura o Paciente .
Medicação semantic_type: @physical_entity	A substância injetada, aplicada, injetada etc. projetada para curar o Paciente .
Paciente	Este é o sofredor da lesão, doença ou dor. No entanto, o Paciente pode não ser expresso como um elemento de frame separado, mas frequentemente ocorre como o possuidor da Aflicção .
Parte_do_corpo semantic_type: @body_part	A Parte_do_corpo é a área específica do corpo do Paciente que é tratada.
Tratamento	É usado para tratamentos e meios, distinguidos por EFs separados.

Figura 33: Frame Cura

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Outros frames já existentes na base de dados da FrameNet Brasil serviram de base para novos frames criados, como, por exemplo, o frame de Artefato, já citado, (que é herdado por Materiais_de_uso_em_saúde ou Prótese) e o frame Substâncias (herdado por Substâncias_corporais), também já mencionado anteriormente.

A metodologia *bottom-up* também retornou em análise alguns termos no *corpus* que evocam frequentemente outros frames, porém estes não podem ser

tratados como específicos do domínio, tais como `Período_de_tempo`, `Cor`, e `Excretar`. A modelagem do domínio se mostra essencial, não só para a análise relativa ao projeto no qual essa dissertação se insere, mas amplificando a base de dados da FrameNet Brasil e conhecimentos relativos ao domínio, visando auxiliar nesta e em outras futuras pesquisas. Na Figura 34, podemos ver todos os frames do domínio com suas respectivas relações, sendo que a natureza das relações podem ser verificadas pelas cores descritas na legenda anexa.

A rede de frames elaborada, até então, para a modelagem deste domínio é composta por vinte e nove frames, com dezessete frames recém-criados e dezesseis frames preexistentes, dos quais seis foram modificados, de forma que atendessem melhor ao domínio. A totalidade de frames incorpora um total de 2322 Unidades Lexicais.

- RELAÇÕES
- Herança
 - Subframe
 - Uso
 - Causativo de
 - Incoativo de
 - Precedência

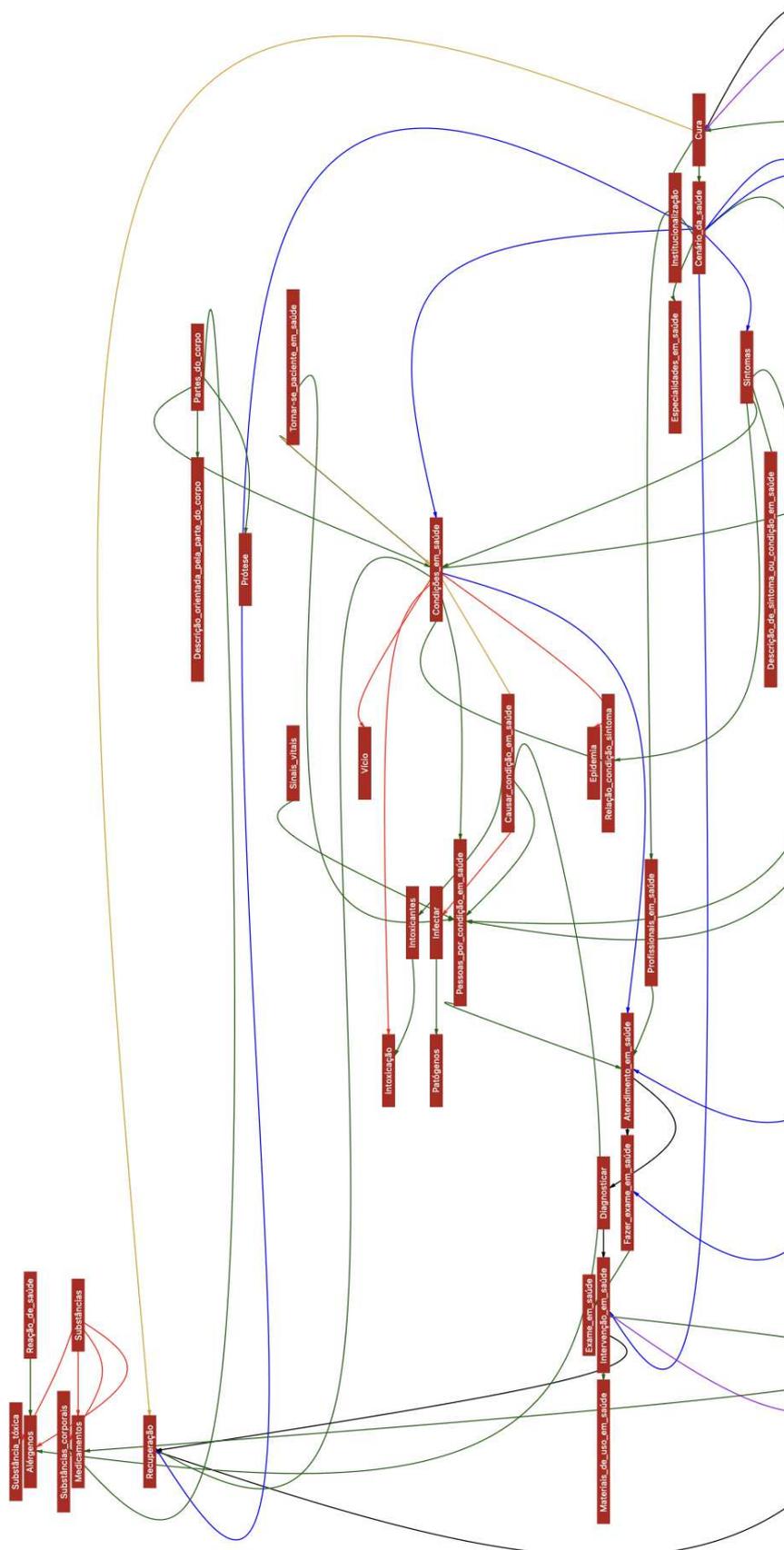


Figura 34: Frames do Domínio da Saúde

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

4.2 FRAMES DO DOMÍNIO DA VIOLÊNCIA

A modelagem de frames relativos ao domínio da violência ocorreu de forma semelhante à modelagem dos frames do domínio da saúde no que concerne às etapas do processo e ao uso das metodologias *top-down* e *bottom-up*. A primeira etapa consistiu em verificar os frames preexistentes relativos ao assunto da violência na base de dados da FrameNet Brasil e averiguar os ajustes necessários, que seriam relativos aos lexemas extraídos dos candidatos a termos obtidos do *corpus* do SINAN, Sistema de Notificação de Agravos Nacional, já mencionado nas seções anteriores. A análise dos frames que já existiam na base de dados e se relacionavam ao domínio consiste na metodologia *top-down*.

A FrameNet Brasil já contava com o frame de Violência, derivado do frame *Violence*, proveniente da FrameNet Berkeley. Na Figura 35, podemos ver o frame originário de Berkeley.

Violence [Lexical Unit Index](#)

Definition:

The words in this frame describe acts (or situations characterized by acts) that cause injury or harm. The acts may involve an **Aggressor** or **Cause** injuring a **Victim** or **Aggressors** causing each other harm.

Research on whether concealed carry laws increase **gun VIOLENCE among gang members** has been inconclusive.

They committed **BRUTALITY** against innocent civilians.

The storm's VIOLENCE last week came without warning and caused unprecedented damage to coastal regions.

FEs:

Core:

Aggressor [agr] The **Aggressor** is the person causing the **Victim's** injury.
 Semantic Type: Sentient **You** **BRUTALITY** on your enemies seems to know no boundaries.

Aggressors [ags] The **Aggressors** commit acts of violence upon each other.
 Hope for a peaceful resolution is gone as **VIOLENCE** between the two sides reignites.

Cause [Cause] The **Cause** marks expressions that indicate some non-intentional, typically non-human, force that inflicts harm on the **Victim**.
 The storm's VIOLENCE left many lives in shambles.

Victim [vic] The **Victim** is the being or entity that is injured.
 Semantic Type: Sentient Increasing **VIOLENCE** against supporters of government reform have attracted international attention.

Figura 35: Frame *Violence*

Fonte: <https://framenet2.icsi.berkeley.edu/>

A Figura 35 apresenta a descrição do frame e os elementos de frame nucleares, que consistem no agressor, na agressão, na vítima e numa causa, como uma possível motivação para a agressão. De maneira semelhante ao procedimento realizado para a modelagem de frames do domínio da saúde, esse frame foi

modificado na FrameNet Brasil para `Cenário_da_Violência` e podemos verificá-lo na Figura 36.

Cenário_da_violência

Definição	
Este frame descreve atos (ou situações caracterizadas por atos) que causam ferimento ou dano.	
Elementos de Frame Nucleares	
FE Core:	
Agressor semantic_type: @sentient	O Agressor é a pessoa que causa dano à Vítima .
Agressores	Os Agressores cometem atos de violência em conjunto.
Causa	A Causa é marcada por expressões que indicam alguma força não intencional, tipicamente não humana que impõe dano à Vítima .
Vítima semantic_type: @sentient	A Vítima é o ser ou entidade que sofre dano.

Figura 36: Frame `Cenário_da_violência`

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main>

O novo frame do `Cenário_da_Violência` *herda* do frame não lexical de `Cenário_de_sujeição`, sendo que este último tem como *subframes* os frames `Cenário_de_Crime` e `Cenário_de_risco`, além de usar o frame `Causar_dano`, que será demonstrado mais adiante. Os frames do `Cenário_de_sujeição`, `Cenário_de_crime` e `Cenário_de_risco` estão presentes nas Figuras 37, 38 e 39, respectivamente.

Cenário_de_sujeição

[@Scenario] [@Generic @Violence] [@Non-lexical @Scenario] [#1403]

Definição	
Nesse frame, ocorre um Evento e uma Entidade é afetada por ele.	
Elementos de Frame Nucleares	
FE Core:	
Entidade	Aquilo que é afetado pelo Evento .
Evento semantic_type: @state_of_affairs	Uma mudança no mundo que afeta a Entidade .
Elementos de Frame Não-Nucleares	
Circunstâncias	Circunstâncias descrevem o estado do mundo (em um determinado momento e local) que é especificamente independente do evento em si e de qualquer um de seus participantes.
Explicação	a Explicação para a qual a Entidade é submetida ao Evento , a Explicação denota uma proposição da qual a cláusula principal (encabeçada pelo alvo) segue logicamente. Isso geralmente significa que a Explicação causa a proposição do alvo, mas não em todos os casos.
Lugar	O local em que a Entidade passa pelo Evento .
Propósito	O Propósito para o qual a Entidade passa pelo Evento .
Tempo	o Tempo em que a Entidade passa pelo Evento .

Figura 37: Frame `Cenário_de_sujeição`

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main>

Cenário_de_crime

[@Scenário] [@Legal] [@Non-lexical @Scenário] [#445

Definição

Um (suposto) **Crime** é cometido e chama a atenção das **Autoridades**. Em resposta, há uma Investigação_criminal e (muitas vezes) a prisão e processos judiciais criminais. A investigação, detenção e outras partes do Processo_criminal são buscadas para encontrar um **Suspeito** (que então pode entrar no Processo_criminal para se tornar o réu) e determinar se esse **Suspeito** corresponde ao **Autor** do **Crime**, e também para determinar se as Acusações correspondem ao **Crime**. Se o **Suspeito** for considerado como tendo cometido o **Crime**, então ele geralmente recebe uma punição proporcional às Acusações.

Elementos de Frame Nucleares

FE Core:

- Acusação** Uma descrição de um tipo de ato que não é permissível de acordo com a lei da sociedade.
- Autor** O indivíduo que comete um **Crime**.
- semantic_type: @sentient**
- Autoridades** O grupo que é responsável pela manutenção da lei e da ordem e, como tal, recebeu o poder de investigar **Crimes**, encontrar **Suspeitos** e determinar se um **Suspeito** deve ser submetido ao Processo_criminal.
- Crime** Um ato, geralmente intencional, que corresponde à descrição que pertence a uma **Acusação** oficial.
- Suspeito** O indivíduo que está sob suspeita de ter cometido o **Crime**.

Figura 38: Frame Cenário_de_crime

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Cenário_de_risco

[@Scenário] [@Generic @Violence] [@Non-lexical @Non-perspectivalized @

Definição

Um **Ativo** está em uma **Situação** específica, que tem a probabilidade de levar ou convidar a um **Evento_nocivo** que afetará negativamente o **Ativo**.

Elementos de Frame Nucleares

FE Core:

- Ativo** Algo julgado desejável ou valioso que pode ser perdido ou danificado.
- Evento_nocivo** Um evento que pode ocorrer ou um estado que pode resultar na perda ou dano do **Ativo**.
- Situação** A **Situação** sob a qual o **Ativo** está seguro ou sob risco.

Elementos de Frame Não-Nucleares

- Grav** Um modificador que expressa o desvio do nível real de segurança do valor esperado dado o **Ativo**, a **Situação** e o estado indicados pelo próprio alvo.
- Lugar** O **Lugar** no qual o grau de segurança se mantém.
- Tempo** O **Tempo** durante o qual o grau de segurança se mantém.

Figura 39: Frame Cenário_de_risco

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Podemos verificar que o frame **Cenário_de_sujeição** funciona como um frame “pai” para os demais cenários. Ele demonstra uma situação mais genérica, na qual um evento afeta uma entidade, sendo que os demais cenários se mostram mais específicos com relação ao domínio da violência e a cada evento em particular, possuindo elementos de frame variados. Ainda relacionado a cenários do domínio da violência, temos o **Cenário_de_tiro**, que *herda* do frame **Cenário_de_tentativa_de_interação_distante**, o qual, por sua vez, é menos relacionado ao domínio que estamos analisando. O frame **Cenário_de_tiro** pode ser visto na Figura 40.

Cenário_de_tiro

[@Scenario] [@Violence @

Definição	
Uma Arma_de_fogo é disparada, causando o movimento de um Projétil . Esse frame existe primariamente para relacionar frames.	
Elementos de Frame Nucleares	
FE Core:	
Agente semantic_type: @sentient	O Agente é a pessoa realizando o ato intencional causando o movimento do Projétil .
Projétil semantic_type: @physical_entity	A entidade cujo movimento rápido e perigoso é causado pelo Agente e/ou Arma_de_fogo .

Figura 40: Frame Cenário_de_tiro

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

Relacionados ao Cenário_de_tiro, devemos citar os frames Portar_armas e Disparar_projéteis, já que também se relacionam com o domínio da Violência, que estamos modelando.

Assim como os frames mencionados acima, apesar de não pertencer diretamente ao domínio da Violência, um frame relacionado, já preexistente na base de dados da FrameNet Brasil, que se mostra também um frame “pai”, de extrema importância para o assunto deste trabalho é o frame Afetar_intencionalmente, apresentado na Figura 41.

Afetar_intencionalmente

[@Action] [@Generic @Violence] [@L

Definição	
Um Agente faz com que um Paciente seja afetado, às vezes por um Meio em particular ou pelo uso de um Instrumento .	
Elementos de Frame Nucleares	
FE Core:	
Agente semantic_type: @sentient	A entidade consciente, geralmente uma pessoa, que realiza o ato intencional que afeta o Paciente .
Paciente	É a entidade que é afetada e que pode, mas não precisa, passar por uma mudança.
FE Core-Unexpressed:	
Evento semantic_type: @state_of_affairs	Identifica o ato que o Agente executa intencionalmente.

Figura 41: Frame Afetar_intencionalmente

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Ele é herdado por diversos frames da base de dados da FrameNet Brasil, incluindo o de *Estupro*, que será descrito mais adiante, e outros frames como *Atacar*, *Prender*, etc.

Outro frame não-exclusivo do domínio, mas que devemos citar, é o frame de *Assistência*, destacando no *corpus* a assistência prestada às vítimas de violência contra a mulher, seja por profissionais de saúde, pela polícia ou pelas autoridades, que evocam frames que veremos mais adiante. O frame *Assistência* está descrito na Figura 42.

Assistência

[@Action][@Generic@Violence][@Lexical][#

Definição	
Um Ajudante beneficia uma Parte_beneficiada tornando possível o culminar de uma Meta que a Parte_beneficiada tem. Uma Entidade_focal que está envolvida em alcançar a Meta poderá substituí-la.	
Elementos de Frame Nucleares	
FE Core:	
Ajudante	O Ajudante performa alguma ação que beneficia a Parte_beneficiada .
Entidade_focal	Identifica uma Entidade_focal envolvida na realização de uma Meta .
Meta	O desejável estado de coisas em que a Parte_beneficiada está envolvida e que é ativado pelo Ajudante .
Parte_beneficiada	A Parte_beneficiada recebe um benefício de uma ação do Ajudante .

Figura 42: Frame *Assistência*

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main#>

Além dos frames base já mencionados e do frame da *Violência*, que foi citado anteriormente e modificado para que abrangesse melhor o domínio, outros frames nucleares no assunto da violência já existiam na base de dados da FrameNet Brasil, tais como: *Abusar*, *Arma*, *Morte*, *Matar* e *Experienciar_Ferimento_Corporal*. A seguir, apresentaremos cada frame preexistente e as respectivas ULs, começando pelo frame *Abusar*, que pode ser verificado na Figura 43.

Abusar

[@Event] [@Generic @Violence] [@Lexical] [#404]

Definição
Um Abusador trata a Vítima de uma maneira cruel e violenta repetidamente, incluindo danos físicos ou forçando a Vítima a se envolver em uma atividade sexual contra sua vontade. A Vítima usualmente não tem o poder de resistir ou lutar. Um Tip de abuso pode ser indicado. Pode também haver um Grav .
Elementos de Frame Nucleares
<p>FE Core:</p> <p>Abusador <small>semantic_type: @sentient</small> O Abusador sujeita a Vítima a repetidos ataques verbais, físicos, emocionais ou sexuais.</p> <p>Vítima <small>semantic_type: @sentient</small> Quem sofre as ações do Abusador.</p>

Figura 43: Frame *Abusar*Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main>

O frame de *Abusar* tem como elementos de frame nucleares o *Abusador* e a *Vítima*. Apesar de podermos associar o lexema *abuso.n* quase sempre diretamente ao abuso sexual, a definição do termo abrange predominantemente qualquer ação de dano repetitivo à vítima, sendo que o frame é evocado por ULs como *abuso.n*, mas também por ULs como *maus tratos.n* e *manipular.v*. O frame *Abusar* herda do frame *Cometer_crime* e também usa o frame *Causar_dano*. Relacionado ao tema do abuso e do estupro, que trataremos mais adiante, por ser um processo relacionado ao procedimento *bottom-up* para a criação de frames, também temos na base de dados da FrameNet Brasil os frames *Sexo* e *Impedir_ou_permitir*, que não pertencem diretamente ao domínio da violência, sendo provenientes da base de dados de Berkeley e de outros domínios previamente estudados, mas que se inserem no domínio estudado devido a ocorrências no *corpus*.

Outro frame central, relativo ao domínio da violência, e que já fazia parte da base de dados da FrameNet Brasil é o frame de *Arma*. Como esperado, ele é evocado por ULs como *revólver.n* e *projétil.n*. No entanto, relativamente ao *corpus* no qual esse estudo se baseia, pudemos verificar que o frame também é evocado por ULs como *pedra.n* e *cano.n*. O frame herda de *Artefato* e é herdado por *Dispositivo_explosivo_improvisado*. Ele é usado pelos frames *Causar_dano* e *Portar_armas*. O frame *Arma* pode ser verificado na Figura 44.

Arma

[@Artifact] [@Violence @Wea

Definição
A Arma é um artefato criado especialmente para causar dano ou prejuízo. Em muitos casos, um Uso , Material de construção, Parte ou Tipos podem ser rigorosamente especificados.
Elementos de Frame Nucleares
<p>FE Core:</p> <p>Arma semantic_type: @artifact A entidade prototipicamente usada pelo Atirador para causar danos.</p>

Figura 44: Frame *Arma*

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main#>

Relacionados ao domínio da violência, evidentemente, frames como *Morte* e *Matar* também se fazem presentes. Apesar da aparente semelhança, esses frames designam cenas diferentes que podem ser demonstradas pelos elementos de frame nucleares e pelo tipo de ULs evocadoras. O frame *Morte*, por exemplo, é evocado por ULs como *óbito.n* e *falecimento.n*, e tem como *causativo* o frame *Matar*, enquanto o frame *Matar* é evocado por ULs como *homicídio.n* e *feminicídio.n*, sendo esta última de óbvia relevância para o este trabalho e, conforme esperado, presente no *corpus* estudado. Os frames *Morte* e *Matar* podem ser vistos nas Figuras 45 e 46.

Morte

Definição
As palavras neste frame descrevem a morte de um Protagonista . A Causa da morte também pode ser expressa.
Elementos de Frame Nucleares
<p>FE Core:</p> <p>Protagonista semantic_type: @sentient É o ser ou entidade que morre.</p>

Figura 45: Frame *Morte*

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main#>

Matar

Definição	
Um Assassino ou uma Causa causa a morte da Vítima .	
Elementos de Frame Nucleares	
FE Core:	
Assassino excludes: Causa	A pessoa ou entidade senciante que causa a morte da Vítima .
Causa excludes: Assassino	Uma entidade inanimada ou processo que causa a morte da Vítima .
Instrumento excludes: Causa semantic_type: @physical_entity	O dispositivo usado pelo Assassino para provocar a morte da Vítima .
Meio excludes: Causa semantic_type: @state_of_affairs	O método ou ação que o Assassino ou a Causa realiza resultando na morte da Vítima .
Vítima semantic_type: @sentient	A entidade viva que morre como resultado do assassinato.

Figura 46: Frame Matar

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Ainda relacionado ao domínio, a base de dados da FrameNet Brasil também já contava com o frame de `Experienciar_ferimento_corporal`, disponível na Figura 47.

Experienciar_ferimento_corporal		[@Event] [@Body @Vio]
Definição		
Um Experienciador sofre um ferimento em uma Parte_do_corpo , apesar de, em alguns casos, nenhuma parte específica ser mencionada. Geralmente, uma Entidade_feridora na qual o Experienciador se machuca é mencionada.		
Elementos de Frame Nucleares		
FE Core:		
Experienciador semantic_type: @sentient	O ser ou entidade que se machuca.	
Parte_do_corpo semantic_type: @body_part	O local do corpo do Experienciador afetado pelo ferimento.	
FE Core set(s): {Experienciador,Parte_do_corpo}		

Figura 47: Frame Experienciar_ferimento_corporal

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

O frame é evocado em diversos domínios, mas se mostra especialmente importante para o domínio da Violência. Situações que envolvem queda ou, por exemplo, corte acidental da própria pessoa enquanto conserta algo ou cozinha,

envolvem a evocação do frame `Experienciar_ferimento_corporal`. Ele é evocado por ULs como: *perfuração.n*, *fraturar.v*, *arranhar.v* e *mutilação.n*.

Tratando-se do domínio da violência, um frame importante, já mencionado, e também já existente na base de dados da FrameNet Brasil é o de `Causar_dano`. Ele abrange exclusivamente as situações de dano físico, envolvendo um Agente (agressor), uma Parte do corpo, uma possível Causa e uma Vítima. Ele é evocado por ULs como *agredir.v*, *ameaçar.v* ou *brigar.v*. A Figura 48 apresenta o frame.

Causar_dano [@Causation] [@Generic @Violence] [@Lexical] [#

Definição

As palavras neste frame descrevem uma situação na qual um **Agente** ou uma **Causa** fere uma **Vítima**. A **Parte_do_corpo** da **Vítima** que é mais diretamente afetada também pode ser mencionada no lugar de **Vítima**. Em tais casos, a **Vítima** é geralmente indicada como um modificador possessivo da **Parte_do_corpo**, no caso em que o EF **Vítima** é indicado em uma segunda camada de EF.

Elementos de Frame Nucleares

FE Core:

Agente
excludes: Causa O **Agente** é a pessoa que causa o ferimento na **Vítima**.
semantic_type: @sentient

Causa
 A **Causa** assinala expressões que indicam alguma força não-intencional, tipicamente não-humana, que causa o ferimento na **Vítima**.

Parte_do_corpo
semantic_type: @body_part A **Parte_do_corpo** identifica o local onde o ferimento aconteceu.

Vítima
semantic_type: @sentient A **Vítima** é o ser ou a entidade que é machucada. Se a **Vítima** estiver incluída na frase que indica a **Parte_do_corpo**, a **Vítima** é etiquetada com EF de segunda camada.

Figura 48: Frame `Causar_dano`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Considerando o cenário, é importante mencionarmos frames como `Correr_risco` ou `Estar_em_risco`. Apesar de aparentarem descrever uma mesma cena, no primeiro frame temos um protagonista, uma entidade que executa uma ação arriscada. No segundo frame, temos um bem de valor que, por força maior, está em situação de risco. A descrição dos frames demonstra a diferença entre o protagonista que se arrisca e o bem que eventualmente está em risco. Podemos verificar as diferentes descrições e elementos de frame nas Figuras 49 e 50.

Correr_risco

[@Action] [@Generic @Violence] [@Lexical] [#348]

Definição

Um **Protagonista** é descrito como sendo exposto a uma situação potencialmente perigosa que pode acabar em um **Resultado_ruim** para si mesmo. Um **Recurso** que está em perigo de perda pode estar no lugar do **Resultado_ruim**. Não há implicação que o **Protagonista** expõe intencionalmente a si mesmo a uma situação arriscada. O **Protagonista** pode estar tentando alcançar algum propósito, que envolve estar em uma situação perigosa. A **Severidade** do risco envolvido também pode ser expressada.

Elementos de Frame Nucleares

FE Core:

- Ação** A **Ação** que cria o risco.
- Protagonista**
semantic_type: @sentient A pessoa que está em risco de um **Resultado_ruim**.
- Recurso**
excludes: Resultado_ruim Algo desejável possuído ou diretamente associado com o **Protagonista** que pode ser perdido ou danificado.
- Resultado_ruim** Uma situação que o **Protagonista** gostaria de evitar.

Figura 49: Frame `Correr_risco`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Estar_em_risco

[@State] [@Generic @Violence] [@Lexical]

Definição

Um **Bem** está em um estado em que é exposto a algo que pode ser afetado por um **Evento_danoso**, o qual pode ser representado metonimicamente por uma **Entidade_perigosa**. Palavras que expressam segurança relativa (isto é, ausência de risco) estão também neste frame.

Elementos de Frame Nucleares

FE Core:

- Bem** Algo considerado desejável ou valioso que pode ser perdido ou danificado.

FE Core-Unexpressed:

- Entidade_perigosa** A entidade concreta ou abstrata que pode vir a causar danos ao **Bem** ou a perda deste, devido a sua participação em um **Evento_danoso**.
- Evento_danoso**
excludes: Entidade_perigosa Uma ação que pode ocorrer ou um estado que pode conter o que pode resultar na perda ou dano do **Bem**.

Figura 50: Frame `Estar_em_risco`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Assim como quem se arrisca e como quem (ou o quê) está eventualmente em risco, temos o frame `Pessoas_por_vocação` que, em outros domínios, normalmente, pode ser evocado por diversos tipos de profissionais. Nesse domínio, a evocação acontece com profissionais relacionados ao combate da violência, como policiais, forças armadas, bombeiros e outros tipos de trabalhadores atuantes na área. Autoridades podem estar envolvidas, inclusive, evocando o frame próprio, `Autoridade`. Entretanto, em determinadas cenas, podemos perceber que uma pessoa ou profissional pode proteger uma pessoa, ou um grupo de pessoas, por exemplo. Essa situação evoca o frame `Proteger`, descrito na Figura 51.

Proteger

Definição	
Alguma Proteção previne um Perigo de danificar um Recurso .	
Elementos de Frame Nucleares	
FE Core:	
Perigo	Uma situação que poderia danificar o Recurso .
Proteção	A pessoa, entidade ou ação que previne o dano a um Recurso .
Recurso	Algo desejável possuído ou diretamente associado com a Proteção que pode ser perdido ou danificado.

Figura 51: Frame *Proteger*

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Outro frame de grande destaque para o domínio, que já existia na base de dados da FrameNet é o de *Encontro_Hostil*, já estudado anteriormente pela FrameNet Brasil (Fonseca, 2015). Ele *herda* elementos do frame *Agir_intencionalmente* e é *herdado* por *Atividade_de_combate*, frame que faz parte do *Cenário_de_Crime*. Ele é evocado por ULs como *batalha.n*, *bullying.n* e *discussão.n*. Podemos verificá-lo na Figura 52.

Encontro_hostil

[@Action] [@Generic @Violence] [@Lexical] [#86

Definição	
Este frame consiste em palavras que descrevem um encontro hostil entre forças opostas (Lado_1 e Lado_2), conceptualizadas coletivamente como Lados sobre um Problema disputado e / ou para atingir uma Finalidade específica.	
Elementos de Frame Nucleares	
FE Core:	
Assunto	Uma questão não resolvida sobre a qual os dois lados de um encontro hostil estão em desacordo.
Finalidade	O resultado desejado do resultado do encontro hostil para o Lado_1 ou para todos os Lados coletivamente. Pode se referir diretamente ao estado de semantic_type : @state_of_affairs coisas desejado, ou o estado de coisas pode estar implícito em uma entidade (abstrata ou concreta) que eles desejam.
Lados semantic_type : @sentient	Os lados expressos em conjunto em um encontro hostil.
Lado_1 requires : Lado_2 excludes : Lados semantic_type : @sentient	Um dos dois participantes em um encontro hostil.
Lado_2 requires : Lado_1 excludes : Lados semantic_type : @sentient	Um dos dois participantes em um encontro hostil, geralmente o segundo mencionado.

Figura 52: Frame *Encontro_hostil*

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Como já relatado anteriormente, o domínio da violência utilizou vários frames já existentes na base da FrameNet Brasil e está relacionado a outros domínios previamente citados, como o `Cenário_de_crime` e o `Cenário_de_tiro`, por exemplo. Relativos a esses cenários ainda temos frames como `Cometer_crime`, `Sequestrar`, `Delitos` e `Roubo`.

Dentre esses, apresentamos, a título de exemplo, o frame de `Cometer_crime`, na Figura 53.

Cometer_crime

Definição	
Um Criminoso (geralmente intencionalmente) comete um Crime , ou seja, faz algo que não é permitido pelas leis da sociedade.	
Elementos de Frame Nucleares	
FE Core:	
Criminoso semantic_type: @sentient	O indivíduo que comete o Crime .
FE Core-Unexpressed:	
Crime	Um ato cometido, geralmente intencional, que é formalmente proibido por lei.

Figura 53: Frame `Cometer_crime`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

O frame é *herdado* por outros que mencionamos logo acima e que serão descritos, como `Roubo`, `Sequestrar` e `Tráfico`, por exemplo, além de ser *subframe* do `Cenário_de_crime`.

Por último, após citarmos, em frames, os mais variados tipos de violência abordados pelo domínio, citaremos o frame que designa a ação de maior importância quando se trata de presenciarmos atos violentos e crimes cometidos, que é o frame `Denunciar`. Sua descrição pode ser vista na Figura 54.

Denunciar

[@Action] [@Generic @Viola

Definição	
Nesse frame, um Informante informa as Autoridades sobre o Comportamento ilegal ou impróprio do Malfeitor .	
Elementos de Frame Nucleares	
FE Core:	
Autoridades	As Autoridades são os oficiais, do governo ou outras, que lidam com a má conduta.
Comportamento	O Comportamento é a ação ou ações pelas quais o Malfeitor é denunciado.
Informante	O Informante denuncia as atividades do Malfeitor às Autoridades .
Malfeitor	O Malfeitor é aquele cujas atividades são denunciadas às Autoridades .

Figura 54: Frame Denunciar

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

O processo de adaptação desses frames para o domínio que está sendo modelado, bem como a distribuição de unidades lexicais presentes no *corpus* estudado para esses frames mais evocados é chamado de *top-down*. Outros frames já existentes na base de dados também foram utilizados, mas, além de representarem o domínio de forma menos relevante para o estudo que está sendo realizado, eles são muito mais numerosos. São exemplos de outros frames evocados pelo domínio: `Decisão_judicial`, `Emergência`, `Restringir_movimento`, etc.

É importante ressaltar que, ainda que os frames apresentados até aqui nesta seção já existissem na base de dados, eles não faziam parte de um cenário de violência. Estavam, portanto, relacionados a outros domínios ou mesmo não relacionados entre si. Também é uma contribuição deste trabalho a organização de tais frames em um cenário de violência.

Relativamente ao procedimento *bottom-up* para modelagem de frames, foram então analisadas as unidades lexicais presentes no *corpus* para que se decidisse quais novos frames deveriam ser criados.

A título de primeira ilustração desse processo, tomemos como exemplo o fato de que a base de dados da FrameNet Brasil já contava com frames como `Sequestrar`, `Roubo`, `Sexo`, no entanto, não havia nenhum frame para lemas que apresentassem a ideia de estupro. Devido à relevância desse conceito para o

domínio, manifesto pela alta frequência relativa de ULs evocadoras dele, e para todo o tema do projeto do qual essa dissertação faz parte, foi criado, então, um frame para *Estupro*. Ele é evocado por ULs que caracterizam o ato em si, como *estuprar.v* e *estupro.n* e é demonstrado na Figura 55.

Estupro [@Action] [@Legal @Violence] [@L

Definição

As palavras neste frame descrevem situações em que o **Perpetrador** tem relações sexuais com a **Vítima** à força (ou ameaçando o uso da força) e sem o seu consentimento.

Elementos de Frame Nucleares

FE Core:

Perpetrador
semantic_type: @sentient É a pessoa que obriga outra pessoa a ter relações sexuais com ela contra a vontade da outra pessoa.

Vítima
semantic_type: @sentient Essa é a pessoa que é forçada a se envolver em relações sexuais contra sua vontade.

FE Core-Unexpressed:

Evento
semantic_type: @state_of_affairs O estupro de uma **Vítima** pelo **Perpetrador**.

Figura 55: Frame *Estupro*

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

O processo de análise do *corpus* permite a percepção de que, identificado um novo frame, os EFs que o compõem podem, por sua vez, ser preenchidos, tipicamente, por ULs que evocam outros frames. No caso do frame de *Estupro*, assim como temos a pessoa num enquadramento criminal, frame já mencionado entres aqueles já existentes na bases de dados da FrameNet, preenchendo o papel de Perpetrador, temos uma possível vítima. Ser vítima, conforme decidido em discussões entre os linguistas participantes do estudo, não é um estado definitivo, assim como outras pessoas em situações vulneráveis também se encontram em uma situação momentânea. Exemplos de situações em que temos a pessoa “vítima”, a pessoa “atingida” ou mencionada de alguma outra forma que esperamos não ser definitiva, tornaram necessária a criação do frame de *Pessoas_por_estado_transitório*. O frame pode ser verificado na Figura 56.

Pessoas_por_estado_transitório

[@Violence] [#1487]

Definição

Este frame contém palavras para indivíduos, conforme vistos em termos de um estado transitório. Eles podem ter um **Descritor**, uma **Origem**, uma **Característica persistente** ou uma **Etnia**. Uma **Idade** específica pode algumas vezes ser especificada, bem como o **Tempo** e o **Lugar**.

Elementos de Frame Nucleares

FE Core:

Estado É um **Estado** temporário e transitório em que uma **Pessoa** se encontra.

Pessoa É um ser humano que se encontra em determinado **Estado**.

Figura 56: Frame `Pessoas_por_estado_transitório`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Assim como temos `Pessoas_por_estado_transitório`, dado o assunto da violência e sendo demonstrados vários itens que integram este cenário, se mostrou também relacionado ao domínio o frame de `Pessoas_por_enquadramento_criminal`. Trata-se de um frame evocado por nomes, normalmente designando um tipo de criminoso envolvido no cenário de violência ou de crime. O frame `Pessoas_por_enquadramento_criminal` pode ser visto na Figura 57.

Pessoas_por_enquadramento_criminal

[@Violence] [#14

Definição

Esse frame contém palavras para indivíduos concebidos em termos de um **Enquadramento_criminal**. A **Pessoa** é concebida como um ser humano descrito. Ela pode ter um **Descritor**, uma **Origem**, uma **Característica persistente** ou uma **Etnia**. Uma **Idade** específica pode, algumas vezes, ser especificada também.

Elementos de Frame Nucleares

FE Core:

Enquadramento_criminal Se relaciona a condição criminal de um indivíduo.

Pessoa A **Pessoa** é o ser humano que possui um **Enquadramento_criminal**.

Figura 57: Frame `Pessoas_por_enquadramento_criminal`

Fonte: <https://webtool.framenetbr.ufjf.br/index.php/webtool/main#>

Um frame evocado por uma UL que delimite uma pessoa que passa por um estado transitório pode englobar diversos tipos de estados não permanentes. Entretanto, verificamos que vários registros apresentados no *corpus* do SINAN mencionavam especificamente pessoas com diversos tipos de ferimentos, como “mordida”, “facada”, “perfurações por projétil”, entre outras. Por isso, mostrou-se necessário criar o frame de `Estar_ferido` especificamente para o domínio. O frame `Estar_ferido` está disponível na Figura 58:

Estar_ferido

Definição	
Um Experienciador se encontra ferido. Com alguma Parte_do_corpo ferida.	
Elementos de Frame Nucleares	
FE Core:	
Experienciador	O ser ou entidade que está ferido.
Parte_do_corpo	O local do corpo do Experienciador afetado pelo ferimento.

Figura 58: Frame `Estar_ferido`

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main>

O *corpus* do SINAN, contém, além das situações de violência, no geral, muitas situações descritivas de vários tipos de acidentes, sejam domésticos, de trânsito, entre outros. Havendo uma quantidade significativa desse tipo de registro, verificou-se a necessidade de criar um frame exclusivo para esse tipo de ocorrência, o frame de `Acidente`. O detalhamento pode ser verificado na Figura 59.

Acidente

Definição	
Este frame é usado para descrever eventos inesperados que podem causar algum tipo de dano.	
Elementos de Frame Nucleares	
FE Core:	
Acidente	O evento inesperado.
Vítima	O indivíduo que sofre o acidente.

Figura 59: Frame `Acidente`

Fonte: <https://webtool.framenetbr.uff.br/index.php/webtool/main>

O frame `Acidente` *herda* do frame `Emergência` e usa o frame `Catástrofe`, via relações frame a frame com frames já existentes na base de dados da FrameNet Brasil. Assim como foi necessária a criação de um frame para

acidentes, verificou-se a partir do *corpus* a necessidade de criação do frame *Descrição_de_acidente*, que é evocado por ULs como *elétrico.n*. Da mesma forma, observando as ocorrências no *corpus*, optamos pela criação do frame *Tráfico*, herdando de *Cometer_crime*.

Ao todo, para o domínio da violência, foram criados cinco novos frames e adicionadas ULs em trinta e oito frames já existentes, totalizando quarenta e três frames e 1768 ULs que integram o domínio. Na Figura 60 podemos ver o Frame Grapher que o ilustra.

Vale ressaltar, por fim, que as novas relações que conectam os 38 frames já existentes entre si, bem como aos novos frames criados, também são contribuições deste trabalho.

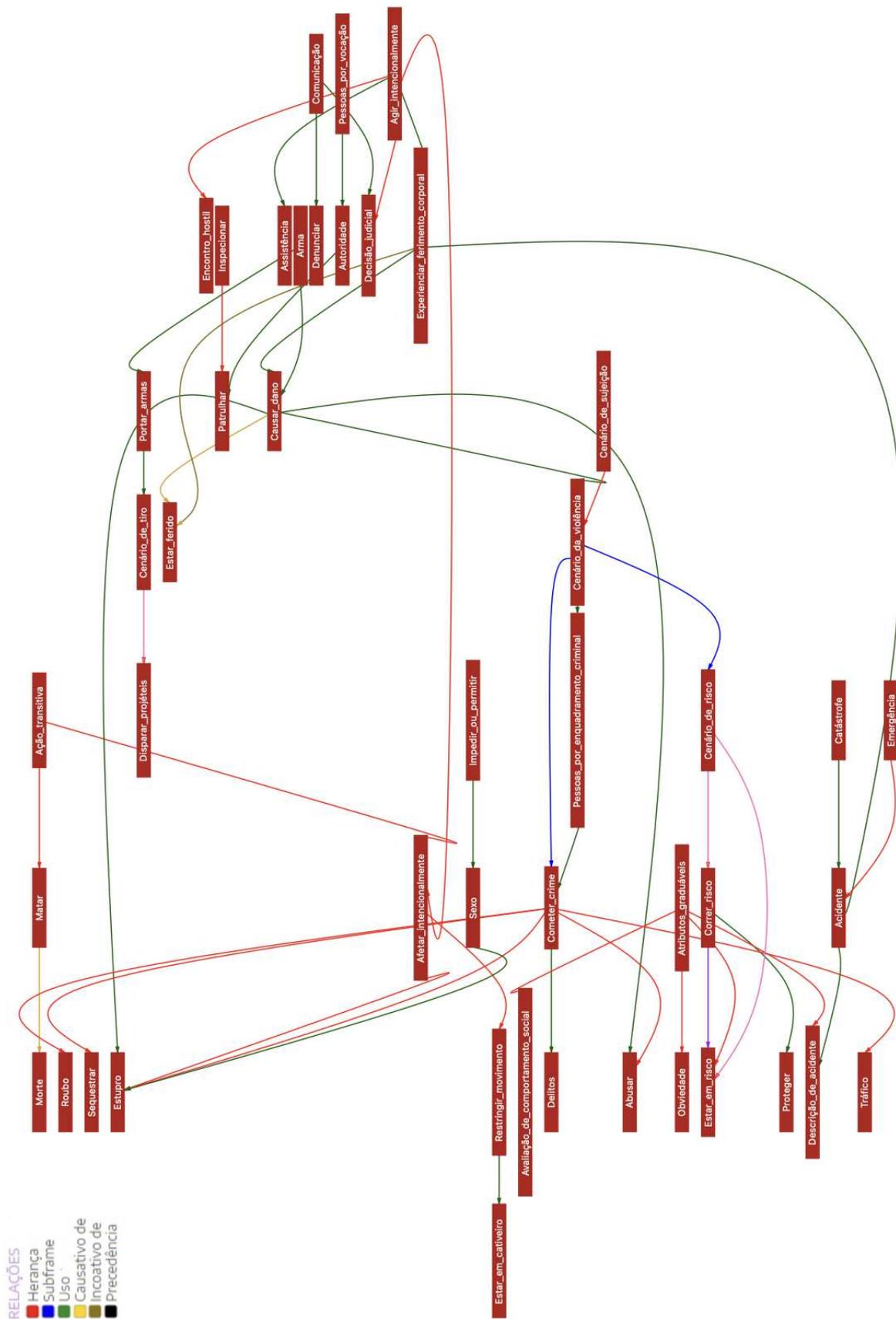


Figura 60: Frame Grapher - Domínio da Violência

Fonte: <https://webtool.frame.net.br/grapher/domain>

5 ANÁLISE DOS FENÔMENOS CAPTURADOS PELO MODELO

Nas últimas seções, descrevemos as motivações do projeto e as necessidades que envolviam a análise dos *corpora* e sua investigação. Em sequência, abordamos a modelagem dos domínios da saúde e da violência, de acordo com as metodologias apresentadas, além da anotação dos *corpora*, descrita na seção 3.5.

A modelagem dos frames, descrita nas seções anteriores, não funciona apenas como mera enumeração das cenas que compõem os domínios da saúde e da violência, mas permite, sobretudo, que possamos identificar padrões com base nas ULs mais evocadoras e nos frames mais evocados por essas ULs. Dessa maneira, foi possível, através da contribuição deste trabalho, um reenquadramento dos frames criados anteriormente via procedimento *top-down*, frames esses que não compunham nenhum domínio em específico, além do reenquadramento de outros frames que estavam contidos em outros domínios.

Para um levantamento de como o léxico da violência se manifesta nos dados dos sistemas públicos de saúde com maior acurácia, foi realizada a anotação automática de todos os *corpora* para frames via LOME (Xia et al., 2021), um sistema multilíngue que implementa um analisador semântico (*parser*), dentre outras possibilidades de processamento de língua natural. A partir das anotações geradas por esse *parser* e com buscas nos dados anotados elaboradas pela equipe de desenvolvimento de sistemas, foi possível gerar um relatório de ULs mais relevantes para os dados das vítimas de violência, bem como os frames que evocam. A partir deste relatório, disponível online¹⁶, apresentamos os léxicos evocadores de ambos os domínios modelados.

5.1 O LÉXICO DA VIOLÊNCIA NOS CAMPOS ABERTOS DOS PRONTUÁRIOS DO E-SUS

Nesta seção, analisaremos o léxico de ambos os domínios modelados a partir das anotações automáticas feitas pelo LOME. Iniciaremos nos debruçando

¹⁶

https://docs.google.com/spreadsheets/d/1kFYV0YVZTb_XOg9OjshqO9RnFPvWoUwyMZpW30lu254/edit?usp=sharing

apenas sobre resultados de anotações feitas no *corpus* de campos abertos dos prontuários do e-SUS, sistema descrito anteriormente na seção 2.1. Apesar de notarmos a presença de muitos frames relacionados diretamente ao domínio da saúde, é importante destacarmos neste *corpus* os atendimentos que ocorreram com mulheres vítimas de violência e, assim, verificarmos os padrões que ocorrem nestes atendimentos. Começando pelo domínio da saúde, iniciaremos o detalhamento dos léxicos evocadores presentes nas anotações feitas nesse *corpus*.

5.1.1 Léxico Evocador dos Frames do Domínio da Saúde

No domínio da saúde, conforme as experiências que tivemos na anotação feita por humanos, ULs que mais evocam frames são termos diretamente associados ao cenário, como *paciente.n*, *gestante.n*, *usuário.n*, *encaminhar.v*, *comprimido.n*, *cabeça.n* etc; evocando frames como *Pessoas_por_condição_em_saúde*, *Condições_em_saúde*, *Sintomas*, *Atendimento_em_saúde* e *Partes_do_corpo*. A UL *paciente.n*, que representa o paciente como substantivo, é a que mais aparece no frame *Pessoas_por_condição_em_saúde*, o que era esperado, devido às várias ocorrências de sentenças como “*Paciente refere...*” no início dos prontuários. Entretanto, ULs que são altamente frequentes em qualquer prontuário – sejam eles de vítimas de violência ou não – são pouco informativas na definição de um padrão lexical da violência de gênero. Assim, a análise que desenvolvemos aqui se debruça sobre as palavras que são relevantes nos prontuários das vítimas.

Tal relevância se traduz em um índice chamado *keyness*, que é medido comparando-se a frequência relativa de frames e ULs no *corpus* geral do e-SUS com a frequência relativa dos mesmos itens no registros de e-SUS que sabemos pertencer a vítimas de violência.

No domínio da saúde, considerando o índice *keyness* do *subcorpus* das vítimas de violência, o frame mais evocado é o de *Intoxicantes*, cuja UL mais evocadora é *chumbinho.n*, levando em conta que este poderia ser ingerido pelas vítimas, em tentativa de suicídio, ou utilizado como veneno contra elas. A frequência do frame *Intoxicantes* se mostra quase três vezes maior para o *subcorpus* de vítimas, com relação ao *corpus* de referência, sendo que ele é também evocado por

ULs como *cigarro.n* e *maconha.n*. A frequência da UL *chumbinho.n* para o *corpus* das vítimas de violência é 19 vezes maior do que para os *corpus* de atendimentos comuns. Outras ULs como *entorpecente.n*, *crack.n*, *maconha.n*, *pedras de crack.n*, *cola.n* e *álcool.n* indicam uso de drogas e entorpecentes relacionado aos casos de violência. A lista de ULs mais evocadoras do domínio da saúde para os casos de vítimas de violência segue a tendência apresentada acima, sendo que o segundo frame mais evocado o de *Vício*, que aparece 2,5 vezes mais do que no *corpus* de referência, seguido do frame *Substância_tóxica*, cujo índice *keyness* aponta ser evocado 1,63 vezes mais. No caso do frame de *Vício*, a UL mais evocadora é *viciado.a*. Considerando a maioria das ocorrências do adjetivo, podemos entender que o fato de a UL constar no *corpus* indica que possuidores de vícios são potenciais perpetradores em casos de violência de gênero, sendo que a ocorrência desta UL é 1,5 vezes maior do que no *corpus* comum. Já para *Substância_tóxica*, a UL mais evocadora é *veneno de rato.n*. Apresentados os dados relacionados a altos índices de relato de uso de entorpecentes relacionados aos casos de Violência de Gênero, com situações de consumo tanto por parte da vítima como por parte do perpetrador, é importante salientar que o uso de entorpecentes por parte da vítima não justifica de nenhuma maneira a agressão sofrida.

A lista prossegue com *Patógenos*, com a UL *tuberculose.n* e *Institucionalização*, com a UL *internar.v* sendo a mais frequente. Apesar de que vários frames do domínio da saúde relacionados às vítimas de violência também dizem respeito ao atendimento em saúde em si, materiais de uso em saúde e ao *Cenário_da_saúde* como um todo, é importante destacarmos as ULs que aparecem no frame de *Medicamentos*. Algumas ULs de alta frequência, como *haldol.n*, *carbonato de lítio.n* e *sertralina.n* indicam uso de medicamentos para tratamento psiquiátrico por parte das vítimas de violência.

Corroboram as análises feitas até aqui as ULs evocadoras do frame de *Condições_em_saúde* com maior *keyness* no *subcorpus* das vítimas de violência. Lidera a lista a UL *intoxicação exógena.n*, 10,37 vezes mais presente nos prontuários de vítimas de violência do que naqueles do *corpus* de referência, seguido das ULs *intoxicação.n*, *intoxicação medicamentosa.n* e *envenenamento.n*, que são evocadas de 9 a 10 vezes mais.

De modo similar, o frame de *Sintomas* indica a situação de vulnerabilidade psicológica das vítimas de violência, tendo como UL mais evocadora o termo *apatia.n*, que aparece quase 7 vezes mais do que em *corpus* comum. Também fazem parte das ULs mais evocadoras para esse frame: *déficit auditivo.n*, *boca seca.n*, *úlceras.n*, *pensamentos suicidas.n*, *desnutrição.n*, *coma.n*, *pesadelo.n*, *ideação suicida.n*, *alucinação.n*, *tremor.v*, *esgotamento.n*, *desânimo.n*, entre outras que podem ser verificadas no relatório.¹⁷ Um dos frames relacionados, *Descrição_de_sintoma_ou_condição_em_saúde*, apresenta a UL *psicológico.a* como mais evocadora e o frame comum *Especialidades_em_saúde* tem como UL mais evocadora o termo *psiquiátrico.a*.

Em suma, os frames mais evocados e ULs mais evocadoras do domínio da saúde no *corpus* de vítimas de violência do e-SUS sugerem uma história de uma paciente em situação de ansiedade, que pode apresentar transtornos psiquiátricos possivelmente decorrentes de atos de violência prévios, de discussões e outras situações de descontentamento. Ela, em muitos casos, se autoflagela, faz uso de medicamentos antipsicóticos, e está em possível contato direto com vícios ou com pessoas viciadas em entorpecentes como álcool e drogas. O léxico evocador dos frames do domínio da violência deverá contar o restante da história, ou seja, os detalhes relacionados às agressões que a vítima acaba vindo a sofrer.

5.1.2 Léxico Evocador dos Frames do Domínio da Violência

Como já relatado na subseção anterior, relativa ao domínio da saúde, os frames mais evocados com as respectivas ULs mais evocadoras do *corpus* dos atendimentos e-SUS vão demonstrar os aspectos do atendimento que muitas vezes são relativos a condições em saúde e sintomas que as mulheres vítimas de violência apresentam previamente a sofrer um ato de violência física.

A manifestação da violência, então, começa já no domínio da saúde, quando olhamos para a análise dos termos que correspondem a sintomas e condições psicológicas que possuem maior índice *keyness*, e se estende para o domínio da

¹⁷https://docs.google.com/spreadsheets/d/1kFYV0YVZTb_XOg9OjshqO9RnFPvWoUwyMZpW30lu254/edit?usp=sharing

violência, quando o índice demonstra a concretização dos atos de violência física e feminicídio.

Para o domínio da violência, ainda em *corpus* e-SUS, temos o frame *Disparar_projéteis* com maior índice *keyness*, ou seja, este frame aparece 6,05 vezes mais no *subcorpus* relativo aos atendimentos realizados com vítimas de violência contra a mulher. A UL mais evocadora desse frame, sendo que ela aparece 21,7 vezes mais do que no *corpus* geral, é o verbo *atirar.v*. O segundo frame mais evocado é o de *Abusar*, com ULs como *assediar.v*, *alienação parental.n*¹⁸, *abuso.n*, *abusar.v*, *ameaçado.a*, *abusivo.a* e *exploração sexual.n*. O frame *Abusar*, como já era esperado, aparece 5,92 vezes mais no *corpus* relacionado a mulheres vítimas de violência do que no *corpus* geral. As ULs apontadas como mais evocadoras do frame *Abusar* aparecem, respectivamente, 13,72 vezes mais para *assediar.v*, 10 vezes mais para *alienação parental.n*, 9 vezes mais para *abuso.n* e *abusar.v*, 6 vezes mais para *ameaçado.a*, 4,4 vezes mais para *abusivo.a* e 1,1 vezes mais para *exploração sexual.n*. O frame de *Cenário_da_violência* aparece em seguida, sendo evocado cinco vezes mais pela UL *violência.n*, sendo que o próximo frame mais evocado em comparação com *corpus* geral é o de *Matar*, com ULs como *homicidar.v*, *suicídio* e *homicídio.n*.

A lista continua com frames como *Pessoa_por_enquadramento_criminal*, sendo evocado 5 vezes mais com a UL *traficante.n*, que aparece 15 vezes mais do que em *corpus* genérico. O frame *Estupro* é evocado 2,4 vezes mais com as ULs *estupro.n* e *estuprar.v* aparecendo 9 e 8 vezes mais respectivamente. O frame *Encontro_hostil* é evocado 2 vezes mais com a UL *batalha.n* sendo a mais evocadora (9,17 vezes mais), entretanto, apesar de aparentar genérico, se apresenta importante, pois apresenta como segunda UL mais evocadora o termo *palavrão.n* (também evocada 9,17 mais), seguida das ULs *enfrentamento.n*, *conflito.n*, *discussão.n* *desentendimento.n* e *briga.n*, que são possíveis pontos de partida para o desencadeamento da violência contra as vítimas.

Apesar de se apresentar como nono frame da lista dos mais evocados no *subcorpus* das vítimas de violência, o frame *Causar_dano* também se mostra

¹⁸ Acredita-se que há equívoco do profissional no registro do termo “alienação” parental que, nos dados, aparenta indicar que a criança sofre negligência. Em exemplo adaptado: “Avó informa que a paciente sofre alienação parental.”

importante, pois nele se encontram ULs que denotam outras formas de violência, além dos danos por projéteis e os auto infligidos. Ele é 1,66 vezes mais evocado no *subcorpus* do que em *corpus* genérico e está na lista por ULs como *apedrejar.v* e *injuriar.v*, aparecendo 17,97 vezes mais, *socar.v*, aparecendo 14,96 vezes mais, *mutilar.v* que surge 10 vezes mais, *violentar.v*, 7,88 vezes mais frequente, *empurrar.v*, 6,48 vezes mais frequente, *agredir.v*, 4,63 mais frequente, *enforcar.v* 4,33 vezes mais frequente, *abandonar.v*, 3,73 vezes mais frequente, entre outras.

O décimo frame mais evocado é de *Arma* (evocado 1,25 vezes mais), surgindo por ULs como *navalha.n* (10,96 vezes mais frequente), *arma branca.n* (5,54 vezes mais frequente), *faca.n* (4,6 vezes mais frequente), *além de projétil.n*, *estilete.n* e *facão.n* sendo 4,63 mais frequentes, dentre outras armas e armas brancas.

Como podemos verificar pelos frames mais evocados e pelas ULs que mais evocam no *corpus*, a violência contra mulher muitas vezes ocorre de maneira extrema, com disparo de projéteis, outras muitas vezes por meio de violência física ou de diversos tipos de armas brancas, sendo que também ocorre na forma de abuso, ameaça, exploração sexual e estupro, por exemplo, podendo ter como fatores pré-determinantes termos já relatados na seção anterior, como ansiedade, transtornos psiquiátricos e abuso de álcool e drogas por ambas as partes.

Resta-nos agora observar como esses atos de violência já reportados pelos profissionais da atenção básica, através do e-SUS, chegam às notificações do SINAN.

5.2 O LÉXICO DA VIOLÊNCIA NOS CAMPOS ABERTOS DO SINAN

Como já relatado na seção 2.1, o SINAN é o Sistema de Informações de Agravos de Notificação, no qual os casos de violência já consumados devem ser registrados por profissionais de saúde. A análise do *subcorpus* proveniente desse sistema apresenta diversas semelhanças com o que já vimos até aqui no e-SUS. Predominantemente, no relacionado ao domínio da violência, vários frames e ULs mais evocados que já vimos anteriormente se manifestam. No entanto, é importante analisarmos e ressaltarmos alguns contrastes entre as anotações feitas para os *subcorpora* dos dados provenientes de ambos os sistemas.

Começando, então, pela análise do domínio da saúde nas anotações feitas em *subcorpus* do SINAN, o frame mais evocado é o de *Causar_condição_em_saúde*, possuindo índice *keyness* 36,35, com a UL *ingerir.v*, presente 344,09 vezes mais do que em *corpus* genérico. Em seguida, temos o frame de *Vício*, evocado 29,2 vezes mais nesse *corpus*. As ULs que mais o evocam são *drogado.n* (substantivo) e *drogado.a* (adjetivo), ambas 73,58 vezes mais frequentes. Depois, temos as ULs *viciado.n* e *viciar.v*, 37,02 e 30,58 vezes mais frequentes, respectivamente. O terceiro frame mais evocado é o de *Substância_tóxica*, com as ULs *veneno de rato.n*, 342,08 vezes mais presente, seguida das ULs *veneno.n*, *veneno de barata.n* e *carrapaticida.n*, com índices *keyness* 335,79, 355,54 e 53, ou seja, a UL *veneno de rato* é encontrada 342,08 vezes mais no *subcorpus* de vítimas de violência do que no *corpus* de referência. A lista segue com o frame *Intoxicantes*, que é evocado 15,65 vezes mais com a UL *chumbinho.n*, que já vimos anteriormente, sendo que, no *corpus* SINAN, ela aparece 620,58 vezes mais do que em *corpus* genérico. A seguir temos as ULs *substância psicoativa.n*, *bala.n*, *bebida alcoólica.n*, *bebida.n*, *crack.n*, *droga.n*, *cerveja.n*, *pedras de crack.n*, *água sanitária.n*, *maconha.n*, *cola.n*, *produto químico.n* e *cocaína.n*, ULs que aparecem de 4,99 a 83,59 vezes mais, indicando tentativa de suicídio por parte das vítimas e uso de entorpecentes pelos envolvidos da violência.

No frame *Institucionalização*, evocado 13,15 vezes mais, temos a UL e *admitir.v*, com índice de 131,41. O quinto frame mais evocado já é o de *Cura*, aparecendo 4,75 mais com as ULs *alta hospitalar.n*, *alta.n* e *alta médica.n*, que aparecem de 20,15 a 21,19 vezes mais, lembrando que é comum que vejamos este frame como um dos mais evocados, porque muitos casos de violência de gênero, diferentemente do que vemos na atenção básica, resultam em internação hospitalar e ocasionam a alta hospitalar quando a vítima, felizmente, não vem a óbito.

O frame *Pessoas_por_condição_em_saúde* é evocado 3,92 vezes mais com as ULs *vítima.n* e *paciente.n* que aparecem até 103,23 vezes mais, no caso da UL *vítima.n*. A lista de frames segue com *Diagnosticar*, *Cenário_da_saúde*, *Relação_condição_sintoma* e *Atendimento_em_saúde* se apresentando como frames comuns relacionados ao domínio e abrangendo ULs como *diagnosticar.v*, *encaminhar.v*, *decorrer.v* e *transferido.a*, respectivamente para cada frame.

Já o frame *Condições_em_saúde*, mesmo sendo muito pouco mais frequente nesse *corpus* – 1,23 vezes –, merece maior atenção pelas ULs que o evocam com mais frequência no *subcorpus* do SINAN, em contraste com os *corpora* de referência. A UL *alcoholizado.a* aparece em primeiro lugar, com índice *keyness* 415,86. Verificando as sentenças relacionadas a essa UL no *corpus*, vemos que o alto consumo de álcool parte tanto da vítima como do agressor, como nos exemplos adaptados: “Paciente alcoholizada apresenta hematomas devido a agressão de parceiro”, ou ainda “Paciente estuprada por homem alcoholizado”. A lista de ULs para o frame *Condições_em_saúde* segue com as ULs: *corte.n*, 153,65 vezes mais evocadora, *envenenamento.n*, 129,38 vezes mais evocadora e *bipolaridade.n*, 64,07, ainda indicando alto índice de tentativas de suicídio por parte das vítimas, sofrimento de agressões através armas brancas e tentativa de envenenamento, como apresenta a sentença adaptada: “Paciente com corte na face, agredida por marido” ou ainda “Paciente com suspeita de evenenamento por rixa familiar”. O frame *Condições_em_saúde* ainda é evocado por ULs como *intoxicação exógena.n*, *desmaiado.a*, *intoxicação medicamentosa.n*, *politraumatismo.n*, *hematoma.n*, *escoriação.n*, *parada cardiorespiratória.n*, *trauma.n*, *internado.a*, *queimadura.n*, *sangramento anal.n*, *sangramento vaginal.n*, entre outras que aparecem até 59,74 vezes mais e sugerem o quadro crítico de gravidade das vítimas que já sofreram estupro ou agressões mais severas.

Assim como o frame *Condições_em_saúde*, o frame *Medicamentos* merece maior detalhamento, sendo evocado 1,17 vezes mais e apresentando ULs como *cartela.n*, que aparece 156,39 vezes mais e *compridos.n*, que aparece 76,72 vezes mais, indicando, mais uma vez, alto índice de tentativa de suicídio por parte das vítimas, sendo ULs encontradas em sentenças como: “Encontrada desacordada ao lado de cartela vazia de diazepam.” A lista segue diferentemente das ULs que evocam o frame de *Medicamentos* no e-SUS, não só com predominância de medicamentos psiquiátricos, mas outros que indicam o tratamento da vítima sob internação. Dessa maneira, ULs como *naldecon.n*, *tilenol.n* (sic), *imosec.n* e *dorflex.n* que são medicamentos para doenças e sintomas comuns, aparecem como mais evocadoras no e-SUS, apesar de ULs como *rivotril.n*, *amplicitil.n*, *antidepressivo.n*, *diazepam.n*, *gardenal.n*, entre outros, que indicam transtornos psiquiátricos, também constarem em grandes números.

Prosseguindo com a lista de frames mais evocados para o domínio da saúde no *subcorpus* do SINAN, temos o frame *Recuperação*, que é evocado por ULs como *medicado.a* e *cicatriz.v*, 7,68 e 1,43 vezes mais presentes, respectivamente, entre outras, que indicam a recuperação das vítimas. O frame seguinte, *Descrição_de_sintoma_ou_condição_em_saúde*, apresenta ULs como *cronicamente.adv*, *convulsivo.a* e *grave.a*, que aparecem 6,02, 5,41 e 5,12 vezes mais, sendo outras ULs mais relacionadas a sintomas psicológicos ou complicações de infecções da vítima sob internação.

Apesar de ser um frame menos evocado diante dos outros já descritos acima, é importante mencionarmos o frame *Partes_do_corpo* com as ULs *cara.n*, 14,36 vezes mais presente, *hemiface.n*, 9,13 vezes e *rosto.n*, 8,8 vezes mais, indicando agressões na face e na região da cabeça, sendo que as ULs mais evocadoras do próximo frame da lista, *Descrição_orientada_pela_parte_do_corpo*: *periorbital.a* e *craniano.a*, com índices *keyness* 22,60 e 20,02, também indicam esse padrão.

O frame de *Sintomas* no *subcorpus* do SINAN para vítimas de violência retrata também os sintomas pós agressão e internação. A lista de ULs se inicia com *desidratar.v*, *convulsionar.v*, *alteração_comportamental.n*, *alteração_dermatológica.n*, *desmaiar.v* e *pensamentos_suicidas.n*, que são de 7,44 a 26,74 vezes mais evocadoras.

Por fim, os frames *Profissionais_em_saúde* e *Intervenção_em_saúde* indicam o tratamento de traumas e outras condições às quais as vítimas foram expostas, via *ortopedia.n*, *neurologia.n*, *internamento.n* (*sic*). As ULs *abortamento.n* e *laparotomia.n*, 9,09 e 5,95 mais frequentes no frame *Intervenção_em_saúde* sugerem a necessidade de aborto e intervenção cirúrgica pós estupro. As ULs mais evocadoras no frame *Especialidades_em_saúde* indicam a necessidade das intervenções e tratamentos psicológicos e psiquiátricos para as vítimas, com as ULs *psicoterapia.n*, 8,5 vezes mais evocadora, e *psiquiatria.n*, 1,06 mais presente, entretanto, a segunda UL mais presente para esse frame é *pediatria.n*, 2,28 vezes mais, indicando que muitas vítimas de violência ainda são menores de idade.

No domínio da violência, os frames mais evocados no *subcorpus* do SINAN retratam a violência sofrida pelas vítimas antes dos atendimentos em saúde ou do

eventual registro de óbito, o qual, por não conter dados linguísticos, não analisamos. A lista de frames mais evocados se inicia com o frame `Pessoas_por_enquadramento_criminal`, evocado 111,85 vezes mais com as ULs *autor.n*, 1198,01 vezes; *agressor.n*, 447,72 mais; *assaltante.n*, 39,72 vezes mais; *acusado.n*, 31,29 vezes mais; entre outras. O segundo frame mais evocado é o de `Disparar_projéteis`, possuindo *keyness* de 54,90. As ULs mais evocadoras para esse frame são: *atirar.v*, com índice *keyness* apontando que ela é 172,15 vezes mais frequente e *tiroteio.n*, com índice de 33,67. Em seguida, o frame `Cenário_da_violência` é o mais evocado, com índice 39,34, sendo mais evocadora a UL *violência.n*, com índice 6,99.

Posteriormente, temos o frame `Matar`, evocado 13,08 vezes mais no *subcorpus* do SINAN para mulheres vítimas de violência. A UL mais evocadora é *homicidiar.v* (sic) com índice 174,95, seguida de *suicidar.v*, com índice 13,24, *matar.v*, com índice 4,4, entre outras de mesmo teor. O frame de `Estupro` é o quinto mais evocado, com índice 9,59, apresentando as ULs *estupro.n* e *estuprar.v*, que são 93,58 e 106,13 mais evocadoras. A lista segue com o frame `Abusar`, que é 8,64 vezes mais evocado. Para esse frame é importante notar que a UL mais evocadora é *alienação parental.n*, evocando 39,72 vezes mais o frame, seguida das ULs *ameaçado.a*, *assediar.v*, *abusar.v*, *abuso.n*, *perseguir.v*, *exploração sexual.n* e *abusivo.a*, que são de 5,27 a 27,26 vezes mais evocadoras. O sexto frame, já mais relacionado ao contexto de violência é o de `Arma` que é evocado 6,94 vezes mais. Suas ULs mais evocadoras são *porrete.n*, 196,16 vezes mais evocadora, *arma branca.n*, 134,17 vezes mais evocadora, *pau.n*, 102,09 vezes mais evocadora, *disparo.n*, 72,12 vezes mais evocadora, *bala perdida.n*, 63,30 vezes mais evocadora, *navalha.n*, 43,2 vezes mais evocadora, *garrafa.n*, 40,01 vezes mais evocadora, *estilete.n*, 36,44 vezes mais evocadora, entre outros tipos de facas que evocam pouco mais de 25 vezes mais. A UL *arma de fogo.n* é 23,80 mais presente no *subcorpus*, sendo que armas brancas ainda são mais utilizadas. No entanto, o alto índice *keyness* da UL *disparo.n* também demonstra que a arma de fogo é muito utilizada.

O sétimo frame mais evocado é o de `Operar_veículo`, evocado 6,5 vezes mais com a UL *pilotar.v*, que é 207,58 mais evocadora, indicando, em sua maioria, eventos de atropelamento. Em seguida, o frame `Delitos` é o mais evocado,

apresentando índice *keyness* 5,22, sendo a UL mais evocadora *negligência.n*, que evoca 5,69 vezes mais. A maioria das sentenças onde encontramos a UL sugere, mais uma vez, violência contra menores de idade, devido à negligência dos pais. O frame *Causar_dano* é o nono frame mais evocado nesse *corpus* e suas ULs mais evocadoras já demonstram a violência consumada. São elas: *abusar.v*, evocando 391,32 vezes mais; *acertar.v*, evocando 196,16 vezes mais; *cotovelar.v*, evocando 196,16 vezes mais; *injuriar.v*, evocando 141,26 vezes mais; *socar.v*, evocando 118,96 vezes mais; *atirar.v*, evocando 108,94 vezes mais; *esmurrar.v*, evocando 104,05 vezes mais; *chutar.v*, evocando 86,93 vezes mais; entre outras de mesmo teor.

O décimo frame mais evocado para o domínio da violência no *subcorpus* do SINAN é o de *Experenciarm_ferimento_corporal*, que é evocado 4,83 vezes mais. A UL mais evocadora para esse frame é *coronhada.n*, sendo 781,64 vezes mais evocadora, seguida da UL *espancamento.n*, que é 372,74 vezes mais frequente; *agredido.a*, 337,83 vezes mais; *soco labial.n* e *soco.n*, 295,39 vezes mais; *chute.n*, 278,49 vezes mais; *garrafada.n*, 216,01 vezes mais; *violentado.a*, 215,13 vezes mais; *pedrada.n*, 189,53 vezes mais; *luta corporal.n*, 172,15 vezes mais; entre outras, indicando em sua maioria o tipo de agressão sofrida pelas vítimas.

A lista de frames segue com *Denunciar* que é evocado 4,40 vezes mais, apresentando ULs como *denunciar.v*, 39,06 vezes mais frequente, *notificar.v*, 26,70 vezes mais, e *encaminhar.v*, 4,41 vezes mais. Apesar de estar na lista de frames mais evocados, o frame poderia estar no topo da lista com maior índice *keyness*, mas ele é apenas o décimo primeiro frame mais evocado, o que pode sugerir que o número de violências denunciadas ainda pode estar abaixo do esperado ou desejável.

As análises de frequência relativa aqui descritas apontam para um perfil de manifestação da violência de gênero nos sistemas públicos de saúde. De um lado, no domínio da saúde, registros nos prontuários do e-SUS demonstram sintomas e condições em saúde que antecedem formas de violência mais graves, de outro lado, as notificações registradas no SINAN já demonstram atos de violência física já consumados, quando as medidas cabíveis podem ser, muitas vezes, apenas paliativas. Na seção 2.1, na qual descrevemos alguns aspectos relativos a violência

de gênero no Brasil e no mundo, verificamos o quanto os números sobre mulheres que sofrem violência já são alarmantes. A anotação de frames dos campos abertos dos sistemas, tanto da atenção primária em saúde como do Sistema SINAN, auxilia na análise dos termos mais frequentes e demonstra o que já foi percebido por outros pesquisadores com relação à subnotificação, alertando para a necessidade de medidas e acompanhamento de sintomas e condições de saúde que aparecem já no âmbito da saúde básica e que podem ser indícios de violências mais graves que já aconteceram ou que virão a acontecer.

6 CONCLUSÃO

Esta dissertação se propôs ser um trabalho que, para além de descrever dois domínios lexicais do português brasileiro a partir da metodologia da FrameNet Brasil, pudesse fazer dessa descrição uma contribuição para um produto tecnológico que busca contribuir para o enfrentamento de um problema social evidente, que é a violência contra as mulheres. Os dados pesquisados sobre o enquadramento da Violência de Gênero no Brasil e no mundo confirmam a necessidade de cada vez mais analisarmos, em várias localidades, registros de atendimentos em saúde, registros de agravos de violência e ainda outros tipos de registros que possam auxiliar em outros estudos e constar em outros sistemas, como ocorrências policiais, entre outros, com devido processo de anonimização para preservar a identidade das vítimas.

A anotação de sentenças baseadas em frames serve para este estudo como um facilitador do entendimento dos dados colhidos por nós pesquisadores e para o processamento e contagem desses dados por máquina via *parser* semântico do sistema LOME, já descrito na seção 5.1, através no qual foi possível obter os índices de frequência dos frames de cada domínio. A análise das frequências de cada frame e respectivas ULs via múltiplas anotações feitas por máquina nos permite identificar padrões relativos às vítimas de violência de gênero, como demonstram as seções 5.1.1 e 5.1.2, de forma que os frames mais evocados e as ULs mais frequentes para cada domínio auxiliam na percepção de quadros de saúde prévios ou concomitantes aos atos de violência registrados, além de apontar para situações que podem ser suspeitas de casos de violência.

O trabalho desenvolvido nesta dissertação contribuiu para o treinamento tanto do modelo de IA do LOME, que rotula semanticamente textos abertos para os frames da Violência e da Saúde, quanto para um segundo algoritmo, em fase finalização, capaz de estimar taxas de subnotificação de violência. No estágio atual do projeto, um painel de visualização de estimativas de subnotificação está sendo construído pela Vital Strategies e pela FrameNet Brasil para ser entregue à prefeitura de Recife, como forma de instrumentalizar gestores em saúde pública com dados que possam direcionar políticas de enfrentamento à violência de gênero, via redução de subnotificação.

REFERÊNCIAS

- ÁVILA, T. & MESQUITA, C. (2020). O Conceito Jurídico de Violência Baseada no Gênero: Um Estudo da Aplicabilidade da Lei Maria da Penha à Violência Fraternal. **Revista Quaestio Juris**, 13(01),174-208.
- AZAMBUJA, M. P. R. & NOGUEIRA, C. (2008). Introdução à Violência Contra as Mulheres como um Problema de Direitos Humanos e de Saúde Pública. **Revista Saúde Soc. São Paulo**, 17(3).
- BAEZA-YATES, R. & NAVARRO, G. (1999). Faster Approximate String Matching . **Algorithmica**, 23, 127–158.
- BOHNER, G. (2001). Writing about rape: use of the passive voice and other distancing text features as an expression of perceived responsibility of the victim. **British Journal of Social Psychology**, 40(4), 515-529.
- BOTT, S.; MORRISON, A.; ELLSBERG, M. (2005). Preventing And Responding To Gender-Based Violence In Middle And Low-Income Countries: A Global Review And Analysis. **World Bank Policy Research Working**, Paper 3618
- BRASIL. (2006). **Lei Maria da Penha**. Lei n. 11.340/2006. Coíbe a violência doméstica e familiar contra a mulher. Brasília: Presidência da República.
- CERQUEIRA, D.; LIMA, R. S. de; BUENO, S. **Atlas da violência 2024**. Rio de Janeiro: IPEA-FBSP, 2024.
- CHOMSKY, N. **Aspects of the Theory of Syntax**. Cambridge, Mass.: MIT. Press, 1965.
- CUNHA, L. F. & RAMALHO, J. C. (2022). NER in Archival Finding Aids: Extended, **Machine Learning and Knowledge Extraction**, 4(1), 42–65.
- DUTRA, L. V. (2024) **Evaluating the Contribution of FrameNet to Gender-Based Violence Identification**. Dissertação de Mestrado em Linguística. University of Gothenburg, GU, Suécia.
- FILLMORE, C. J. (1968). The case for case. In: Bach, E. & Harms, E. (Eds.). **Universals in linguistic theory** (pp.1–68). New York: Holt, Rinehart and Winston.
- FILLMORE, C. J. (1971). Verbs of judging: an exercise in semantic description. In: Fillmore, C. J. & Langendoen, D. T. (Ed.). **Studies in Linguistic Semantics**. New York: Holt, Rinehart and Winston.
- FILLMORE, C. J. (1975). An alternative to checklist theories of meaning. In: **Proceedings of the First Annual Meeting of the Berkeley Linguistics Society**. Berkeley: Berkeley Linguistics Society, pp. 123–131.

FILLMORE, C. J. (1977). Scenes and frames semantics. In: Zampolli, A. (Ed.). **Linguistic Structures Processing: Fundamental Studies in Computer Science**, nº 59. Amsterdam: North Holland Publishing.

FILLMORE, C. J. (1982). Frame semantics. In: Linguistic Society of Korea (Ed.), **"Linguistics in The Morning Calm"**. Seoul: Hanshin, p.111-138.

FILLMORE, C. J. (1985) Frames and the semantics of understanding. **Quaderni di Semantica**, 6(2), 222-254.

FILLMORE, C. J. (2012). Encounters with Language. **Computational Linguistics**, 38(4), 701-718.

FILLMORE, C. J., BAKER, C. (2009). A frames approach to semantic analysis. In Heine, B. & Narrog, H. (Eds). **The Oxford Handbook of Linguistic Analysis**. Oxford: Oxford University Press.

FILLMORE, C. J., PETRUCK, M. R. L., RUPPENHOFER, J. & WRIGHT, A. (2003). FrameNet in Action: The Case of Attaching. **International Journal of Lexicography** 16(3), 297–332.

GAMONAL, M. A. (2013). **Copa 2014 FrameNet Brasil: Diretrizes para a Constituição de um Dicionário Eletrônico Trilíngue a partir da Análise de Frames da Experiência Turística**. Dissertação de Mestrado em Linguística. Universidade Federal de Juiz de Fora.

GARBIN, C. A. S.; DIAS, I. A.; ROVIDA, T. A. S.; GARBIN, A. J. I. (2015). Desafios do profissional de saúde na notificação da violência: obrigatoriedade, efetivação e encaminhamento. **Ciência & Saúde Coletiva**, 20, 1879-1890.

GEERAERTS, Dirk. (2006). **Cognitive Linguistics: basic readings**. Berlin: Mouton De Gruyter.

GRAY, K. & WEGNER, D. M. (2009). Moral typecasting: divergent perceptions of moral agents and moral patients. **Journal of Personality and Social Psychology**, 96(3), 505-520.

GRUBER, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? **International Journal of Human Computation Studies** 43, 907–928.

HUANG, C.-R., CALZOLARI, N., GANGEMI, A., LENCI, A., OLTRAMARI, A., & PRÉVOT, L. (2010). **Ontology and the Lexicon**. Cambridge, Mass.: Cambridge University Press.

HUTTENLOCHER, J.; EISENBERG, K. & STRAUSS, S. (1968). Comprehension: Relation between Perceived Actor and Logical Subject. **Journal of Verbal Learning and Verbal Behavior**, 7, 527-530.

KIND, L.; ORSINI, M. L. P.; NEPOMUCENO, V.; GONÇALVES, L.; SOUZA, G. A.; FERREIRA, M. F. F. (2013). Subnotificação e (in) visibilidade da violência contra mulheres na atenção primária à saúde. **Cadernos de Saúde Pública**, 29, 1805-1815.

LOPES, C. B. (2005). Direitos humanos das mulheres: dois passos à frente, um passo atrás. In: RODRIGUES, A. M. et al. (Org.). **Direitos humanos das mulheres**. Coimbra: Coimbra Ed. p. 157-170.

SINALO, C. W. & MANDOLINI, N. (2023). **Representing Gender-Based Violence - Global Perspectives**. Braga: Palgrave Macmillan.

MELUZZI, C., PINELLI, E., VALVASON, E. & ZANCHI, C. (2021). Responsibility attribution in gender-based domestic violence: A study bridging corpus-assisted discourse analysis and readers' perception. **Journal of Pragmatics**, 185, 73-92.

MINNEMA, G., GEMELLI, S., ZANCHI, C., PATTI, V., CASELLI, T. & NISSIM, M. (2021). Frame Semantics for Social NLP in Italian: Analyzing Responsibility Framing in Femicide News Reports. In **Proceedings of the Eighth Italian Conference on Computational Linguistics CliC-it 2021**. Turin: Accademia University Press, p. 235-242.

MINNEMA, G., GEMELLI, S., ZANCHI, C., CASELLI, T. & NISSIM, M. (2022a). Dead or Murdered? Predicting Responsibility Perception in Femicide News Reports. In **Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long papers)**. Online: Association for Computational Linguistics, p. 1078-1090.

MINNEMA, G., GEMELLI, S., ZANCHI, C., CASELLI, T. & NISSIM, M. (2022b). SocioFillmore: A Tool for Discovering Perspectives. In **Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations**. Dublin: Association for Computational Linguistics, p. 240-250.

PUSTEJOVSKY, J. (1995). **The generative lexicon**. Cambridge, Mass.: The MIT press.

PUSTEJOVSKY, J. & JEZEK, E. (2016). **Integrating generative lexicon and lexical semantic resources**. LREC 2016 Tutorial.

PUSTEJOVSKY, J., & STUBBS, A. (2012). **Natural Language Annotation for Machine Learning: A guide to corpus-building for applications**. New York: O'Reilly Media, Inc.

RUPPENHOFER, J.; ELLSWORTH, M.; PETRUCK, M. R. L.; JOHNSON, C. R.; SCHEFFEZYK, J. (2016) **Framenet II: Extended Theory And Practice**. Berkeley: ICSI, Disponível em: <https://Framenet2.lcsi.Berkeley.Edu/Docs/R1.7/Book.Pdf>.

SAFFIOTI, H. (2015). **Gênero Patriarcado Violência**. São Paulo: Expressão Popular.

SANCHEZ, A. (1995). Definicion e historia de los corpus. In: SANCHEZ, A. et al (org.). **CUMBRE Corpus Linguistico de Espanol Contemporaneo**. Madrid: SGEL.

SARDINHA, T. B. (2004). **Linguística de Corpus**. São Paulo: Manole.

SILVA, G. F. (2009). Semântica de Frames. **Cadernos de Tradução**, 25, 25-54.

SIMISTER, J. (2012). **Gender Based Violence - Causes and Remedies**. New York: Nova Science Publishers Inc.

SOUZA, F.; NOGUEIRA, R. & LOTUFO, R. (2020). BERTimbau: pretrained BERT models for Brazilian Portuguese. In: **9th Brazilian Conference on Intelligent Systems, BRACIS**. Rio Grande do Sul: SBC.

TEDESCHI, S.; MAIORCA, V.; CAMPOLUNGO, N.; CECCONI, F. & NAVIGLI, R. (2021). WikiNEuRal: Combined Neural and Knowledge-based Silver Data Creation for Multilingual NER. In: **Findings of the Association for Computational Linguistics: EMNLP 2021**. Punta Cana: Association for Computational Linguistics, pp. 2521–2533.

TERRY, G. (2007). **Gender-Based Violence**. Oxford: Oxfam GB.

TESNIÈRE, Lucien. (1959). **Éléments de Syntaxe Structurale**. Paris: Klincksieck.

TORRENT, T. T.; MATOS, E.; LAGE, L.; LAVIOLA, A.; TAVARES, T.; ALMEIDA, V. G.; SIGILIANO, N. (2018). Towards continuity between the lexicon and the constructicon in FrameNet Brasil. In: LYNDFELT, B.; BORIN, L.; OHARA, K. H.; TORRENT, T. T. (Eds.). **Constructicography: Constructicon development across languages**. Amsterdam: John Benjamins Publishing Company.

TORRENT, T. T.; MATOS, E. E. S.; BELCAVELLO, F.; VIRIDIANO, M.; GAMONAL, M. A.; COSTA, A. D.; MARIM, M. C. (2022). Representing Context in FrameNet: A Multi-Dimensional, Multimodal Approach. **Frontiers in Psychology**, v. 13, article 838441.

XIA, P.; QIN, G.; VASHSHTHA, S.; CHEN, Y.; CHEN, T.; MAY, C.; HARMAN, C.; RAWLINS, K.; WHITE, A. S.; DURME, B. V. (2021). LOME - Large Ontology Multilingual Extraction. In **Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations**. Online: ACL Anthology, pp. 149–159

ZHOU, K., SMITH, A. & LEE, L. (2021). Assessing Cognitive Linguistic Influences in the Assignment of Blame. In **Proceedings of the Ninth International Workshop on**

Natural Language Processing for Social Media. Online: Association for Computational Linguistics, pp. 61-69.

ANEXOS – DICIONÁRIOS DE DADOS DOS SISTEMAS E-SUS APS, SINAN E SIM



Dicionário de dados

AbastecimentoDeAgua

Código	Descrição	Observações
117	Rede encanada até o domicílio	-
118	Poço / Nascente no domicílio	-
119	Cisterna	-
120	Carro pipa	-
121	Outro	-

AcessoHigiene

Código	Descrição	Observações
42	Banho	-
43	Acesso a sanitário	-
44	Higiene bucal	-
45	Outros	-

AleitamentoMaterno

Código	Descrição	Observações
1	Exclusivo	-
2	Predominante	-
3	Complementado	-
4	Inexistente	-

AnimalNoDomicilio

Código	Descrição	Observações
128	Gato	-
129	Cachorro	-

Código	Descrição	Observações
130	Pássaro	-
131	De criação (porco, galinha...)	-
132	Outros	-

AtencaoDomiciliarOrigem

Código	Descrição	Observações
1	UBS	-
6	Outros	-
11	Hospital	-
12	Unidade pronto atendimento	-
13	CACON / UNACON	-
14	Urgência / Emergência hospital SOS	-
15	Hospital SOS demais setores	-

ConclusaoDestinoElegivel

Código	Descrição	Observações
1	Admissão própria EMAD	-
2	Encaminhado para outra EMAD	-
3	Encaminhado para atenção básica	-
4	Outro encaminhamento	-

ConclusaoDestinoInelegivel

Código	Descrição	Observações
1	Instabilidade clínica	-
2	Necessidade propedeutica	-
3	Outro motivo	-
4	Ausência de cuidador	-
5	Outras condições sociais	-

CondicaoDePosseEUseDaTerra

Código	Descrição	Observações
101	Proprietário	-
102	Parceiro(a) / Meeiro(a)	-
103	Assentado(a)	-
104	Posseiro	-
105	Arrendatário(a)	-
106	Comodatário(a)	-
107	Beneficiário(a) do banco da terra	-
108	Não se aplica	-

CondutaDesfecho

Código	Descrição	Observações
3	Alta administrativa da AD	-
1	Alta clínica da AD	-
2	Encaminhamento para Atenção Básica (AD1)	-
4	Urgência / Emergência	-
5	Internação hospitalar	-
6	Saída por óbito / final de acompanhamento pós-óbito	-

CondutaEncaminhamento

Código	Descrição	Observações
1	Retorno para consulta agendada	-
2	Retorno para cuidado continuado / programado	-
12	Agendamento para grupos	-
3	Agendamento para NASF	-
9	Alta do episódio	-
11	Encaminhamento interno no dia	-
4	Encaminhamento para serviço especializado	-

Código	Descrição	Observações
5	Encaminhamento para CAPS	-
6	Encaminhamento para internação hospitalar	-
7	Encaminhamento para urgência	-
8	Encaminhamento para serviço de atenção domiciliar	-
10	Encaminhamento intersetorial	-

Conduta Encaminhamento Odonto

Código	Descrição	Observações
16	Retorno para consulta agendada	-
12	Agendamento para outros profissionais AB	-
13	Agendamento para NASF	-
14	Agendamento para grupos	-
15	Tratamento concluído	-
1	Atendimento à pacientes com necessidades especiais	-
2	Cirurgia BMF	-
3	Endodontia	-
4	Estomatologia	-
5	Implantodontia	-
6	Odontopediatria	-
7	Ortodontia / Ortopedia	-
8	Periodontia	-
9	Prótese dentária	-
10	Radiologia	-
11	Outros	-

Consideracao Peso

Código	Descrição	Observações
21	Abaixo do peso	-
22	Peso adequado	-

Código	Descrição	Observações
23	Acima do peso	-

CuidadorCidado

Código	Descrição	Observações
1	Não possui	-
2	Cônjuge / Companheiro	-
3	Filho / Enteado	-
4	Pai / Mãe	-
5	Avô(ó)	-
6	Neto	-
7	Irmão	-
8	Outro	-

CursoMaisElevado

Código	Descrição	Observações
51	Creche	-
52	Pré-escola (exceto CA)	-
53	Classe alfabetizada - CA	-
54	Ensino fundamental 1ª a 4ª séries	-
55	Ensino fundamental 5ª a 8ª séries	-
56	Ensino fundamental completo	-
61	Ensino fundamental especial	-
58	Ensino fundamental EJA - séries iniciais (supletivo 1ª a 4ª)	-
59	Ensino fundamental EJA - séries finais (supletivo 5ª a 8ª)	-
60	Ensino médio, médio 2º ciclo (científico, técnico e etc)	-
57	Ensino médio especial	-
62	Ensino médio EJA (supletivo)	-
63	Superior, aperfeiçoamento, especialização, mestrado, doutorado	-
64	Alfabetização para adultos (Mobral, etc)	-

Código	Descrição	Observações
65	Nenhum	-

DeficienciaCidadao

Código	Descrição	Observações
12	Auditiva	-
13	Visual	-
14	Intelectual / Cognitiva	-
15	Física	-
16	Outra	-

Desfecho

Código	Descrição	Observações
1	Visita realizada	-
2	Visita recusada	-
3	Ausente	-

DestinoDoLixo

Código	Descrição	Observações
93	Coletado	-
94	Queimado / Enterrado	-
95	Céu aberto	-
96	Outro	-

DoencaCardiaca

Código	Descrição	Observações
24	Insuficiência cardíaca	-
25	Outro	-
26	Não sabe	-

DoençaRespiratoria

Código	Descrição	Observações
30	Asma	-
31	DPOC / Enfisema	-
32	Outro	-
33	Não sabe	-

FormaDeEscoamentoDoBanheiroOuSanitario

Código	Descrição	Observações
122	Rede coletora de esgoto / pluvial	-
123	Fossa séptica	-
124	Fossa rudimentar	-
125	Direto para um rio / lago / mar	-
126	Céu aberto	-
127	Outra forma	-

FornecimentoOdonto

Código	Descrição	Observações
1	Escova dental	-
2	Creme dental	-
3	Fio dental	-

LocalDeAtendimento

Código	Descrição	Observações
1	UBS	-
2	Unidade móvel	-
3	Rua	-
4	Domicílio	-
5	Escola / Creche	-

Código	Descrição	Observações
6	Outros	-
7	Polo (academia da saúde)	-
8	Instituição / Abrigo	-
9	Unidade prisional ou congêneres	-
10	Unidade socioeducativa	-
11	Hospital	Utilizado apenas na ficha de atendimento domiciliar
12	Unidade de pronto atendimento	Utilizado apenas na ficha de atendimento domiciliar
13	CACON / UNACON	Utilizado apenas na ficha de atendimento domiciliar
14	Hospital SOS Urgência / Emergência	Utilizado apenas na ficha de atendimento domiciliar
15	Hospital SOS demais setores	Utilizado apenas na ficha de atendimento domiciliar

LocalizacaoDaMoradia

Código	Descrição	Observações
83	Urbana	-
84	Rural	-

MaterialPredominanteNaConstrucao

Código	Descrição	Observações
109	Alvenaria com revestimento	-
110	Alvenaria sem revestimento	-
111	Taipa com revestimento	-
112	Taipa sem revestimento	-
113	Madeira emparelhada	-
114	Material aproveitado	-
115	Palha	-
116	Outro material	-

ModalidadeAD

Código	Descrição	Observações
1	AD 1	-
2	AD 2	-
3	AD 3	-
4	Inelegível	Utilizado apenas na ficha de avaliação de elegibilidade

MotivoSaida

Código	Descrição	Observações
135	Óbito	-
136	Mudança de território	-

MotivoVisita

Código	Descrição	Observações
1	Cadastramento / Atualização	#TIPO_VISITA
29	Visita periódica	-
2	Consulta	#BUSCA_ATIVA
3	Exame	-
4	Vacina	-
30	Condicionalidades do bolsa família	-
5	Gestante	#ACOMPANHAMENTO
6	Puérpera	-
7	Recém-nascido	-
8	Criança	-
9	Pessoa com desnutrição	-
10	Pessoa em reabilitação ou com deficiência	-
11	Pessoa com hipertensão	-
12	Pessoa com diabetes	-
13	Pessoa com asma	-

Código	Descrição	Observações
14	Pessoa com DPOC / enfisema	-
15	Pessoa com câncer	-
16	Pessoa com outras doenças crônicas	-
17	Pessoa com hanseníase	-
18	Pessoa com tuberculose	-
32	Sintomáticos respiratórios	-
33	Tabagista	-
19	Domiciliados / Acamados	-
20	Condições de vulnerabilidade social	-
21	Condicionalidades do bolsa família	-
22	Saúde mental	-
23	Usuário de álcool	-
24	Usuário de outras drogas	-
25	Egresso de internação	#OUTROS
26	Controle de ambientes / vetores	-
27	Convite para atividades coletivas / campanha de saúde	-
31	Orientação / Prevenção	-
28	Outros	-

Nacionalidade

Código	Descrição	Observações
1	Brasileira	-
2	Naturalizado	-
3	Estrangeiro	-

NASF

Código	Descrição	Observações
1	Avaliação / Diagnóstico	-
2	Procedimentos clínicos terapêuticos	-

Código	Descrição	Observações
3	Prescrição terapêutica	-

OrientacaoSexual

Código	Descrição	Observações
5	Heterossexual	-
6	Gay	-
7	Lésbica	-
8	Bissexual	-
9	Travesti	-
10	Transsexual	-
11	Outro	-

OrigemAlimentacao

Código	Descrição	Observações
37	Restaurante popular	-
38	Doação grupo religioso	-
39	Doação restaurante	-
40	Doação popular	-
41	Outros	-

PraticasIntegrativasComplementares

Código	Descrição	Observações
1	Medicina tradicional chinesa	-
2	Antroposofia aplicado a saúde	-
3	Homeopatia	-
4	Fitoterapia	-
5	Termalismo / Crenoterapia	-
6	Práticas corporais e mentais em PICS	-

Código	Descrição	Observações
7	Técnicas manuais em PICS	-
8	Outros	-

PraticasTemasParaSaude

Código	Descrição	Observações
1	Alimentação saudável	-
2	Aplicação tópica de fluor	-
3	Saúde ocular	-
4	Autocuidado de pessoas com doenças crônicas	-
5	Cidadania e direitos humanos	-
6	Saúde do trabalhador	-
7	Dependência química / tabaco / álcool / outras drogas	-
8	Envelhecimento / Climatério / Andropausa / Etc	-
9	Escovação dentária supervisionada	-
10	Plantas medicinais / Fitoterapia	-
11	Práticas corporais / Atividade física	-
12	Práticas corporais e mentais em PIC	-
13	Prevenção da violência e promoção da cultura da paz	-
14	Saúde ambiental	-
15	Saúde bucal	-
16	Saúde mental	-
17	Saúde sexual e reprodutiva	-
18	Semana saúde na escola	-
19	Agravos negligenciados	-
20	Antropometria	-
21	Outros	-
22	Saúde auditiva	-
23	Desenvolvimento da linguagem	-
24	Verificação da situação vacinal	-

Código	Descrição	Observações
25	PNCT 1	-
26	PNCT 2	-
27	PNCT 3	-
28	PNCT 4	-

ProblemaRins

Código	Descrição	Observações
27	Insuficiência renal	-
28	Outro	-
29	Não sabe	-

PublicoAlvo

Código	Descrição	Observações
1	Comunidade em geral	-
2	Criança de 0 a 3 anos	-
3	Criança de 4 a 5 anos	-
4	Criança de 6 a 11 anos	-
5	Adolescente	-
6	Mulher	-
7	Gestante	-
8	Homem	-
9	Familiares	-
10	Idoso	-
12	Pessoas com doenças crônicas	-
13	Usuário de tabaco	-
14	Usuário de álcool	-
15	Usuário de outras drogas	-
16	Portador com sofrimento ou transtorno mental	-
17	Profissional da educação	-

Código	Descrição	Observações
18	Outros	-

QuantasVezeAlimentacao

Código	Descrição	Observações
34	1 vez	-
35	2 ou 3 vezes	-
36	mais de 3 vezes	-

RacaCor

Código	Descrição	Observações
1	Branca	-
2	Preta	-
4	Parda	-
3	Amarela	-
5	Indígena	-

RelacaoParentesco

Código	Descrição	Observações
137	Cônjuge / Companheiro(a)	-
138	Filho(a)	-
139	Enteado(a)	-
140	Neto(a) / Bisneto(a)	-
141	Pai / Mãe	-
142	Sogro(a)	-
143	Irmão / Irmã	-
144	Genro / Nora	-
145	Outro parente	-
146	Não parente	-

RendaFamiliar

Código	Descrição	Observações
1	1/4 de salário mínimo	-
2	Meio salário mínimo	-
3	Um salário mínimo	-
4	Dois salários mínimos	-
7	Três salários mínimos	-
5	Quatro salários mínimos	-
6	Acima de quatro salários mínimos	-

ResponsavelCrianca

Código	Descrição	Observações
1	Adulto responsável	-
2	Outra(s) criança(s)	-
133	Adolescente	-
3	Sozinha	-
134	Creche	-
4	Outro	-

Sexo

Código	Descrição	Observações
0	Masculino	-
1	Feminino	-

SituacaoDeMoradia

Código	Descrição	Observações
75	Próprio	-
76	Financiado	-
77	Alugado	-

Código	Descrição	Observações
78	Arrendado	-
79	Cedido	-
80	Ocupação	-
81	Situação de rua	-
82	Outra	-

SituacaoExame

Código	Descrição	Observações
"S"	Solicitado	-
"A"	Avaliado	-

SituacaoMercadoTrabalho

Código	Descrição	Observações
66	Empregador	-
67	Assalariado com carteira de trabalho	-
68	Assalariado sem carteira de trabalho	-
69	Autônomo com previdência social	-
70	Autônomo sem previdência social	-
71	Aposentado / Pensionista	-
72	Desempregado	-
73	Não trabalha	-
74	Outro	-

TemasParaReuniao

Código	Descrição	Observações
1	Questões administrativas / funcionamento	-
2	Processo de trabalho	-
3	Diagnóstico / Monitoramento do território	-

Código	Descrição	Observações
4	Planejamento / Monitoramento das ações da equipe	-
5	Discussão de caso ou projeto terapêutico singular	-
6	Educação permanente	-
7	Outros	-

TempoSituacaoDeRua

Código	Descrição	Observações
17	Menos de 6 meses	-
18	6 a 12 meses	-
19	1 a 5 anos	-
20	Mais de 5 anos	-

TipoAtividadeColetiva

Código	Descrição	Observações
1	Reunião de equipe	-
2	Reunião com outras equipes de saúde	-
3	Reunião intersetorial / Conselho local de saúde / Controle social	-
4	Educação em saúde	-
5	Atendimento em grupo	-
6	Avaliação / Procedimento coletivo	-
7	Mobilização social	-

TipoDadoSerialized

Código	Descrição	Observações
2	Ficha de Cadastro Individual	-
3	Ficha de Cadastro Domiciliar	-
4	Ficha de Atendimento Individual	-
5	Ficha de Atendimento Odontológico	-

Código	Descrição	Observações
6	Ficha de Atividade Coletiva	-
7	Ficha de Procedimentos	-
8	Ficha de Visita Domiciliar	-
-	-	-
10	Ficha de Atendimento Domiciliar	-
11	Ficha de Avaliação de Elegibilidade	-
12	Marcadores de Consumo Alimentar	-

TipoDeAcessoAoDomicilio

Código	Descrição	Observações
89	Pavimento	-
90	Chão batido	-
91	Fluvial	-
92	Outro	-

TipoDeAtendimento

Código	Descrição	Observações
1	Consulta agendada programada / Cuidado continuado	-
2	Consulta agendada	-
4	Escuta inicial / Orientação	-
5	Consulta no dia	-
6	Atendimento de urgência	-
7	Atendimento programado	Utilizados apenas nas fichas de atendimento domiciliar
8	Atendimento não programado	Utilizados apenas nas fichas de atendimento domiciliar

TipoDeConsultaOdonto

Código	Descrição	Observações
1	Primeira consulta odontológica programática	-

Código	Descrição	Observações
2	Consulta de retorno	-
4	Consulta de manutenção	-

TipoDeDomicilio

Código	Descrição	Observações
85	Casa	-
86	Apartamento	-
87	Cômodo	-
88	Outro	-

TratamentoDeAguaNoDomicilio

Código	Descrição	Observações
97	Filtração	-
98	Fervura	-
99	Cloração	-
100	Sem tratamento	-

Turno

Código	Descrição	Observações
1	Matutino	-
2	Vespertino	-
3	Noturno	-

VigilanciaEmSaudeBucal

Código	Descrição	Observações
1	Abscesso dento alveolar	-
2	Alteração em tecidos moles	-
3	Dor de dente	-

Código	Descrição	Observações
4	Fendas / Fissuras lábio palatais	-
5	Fluorose dentária moderada / severa	-
6	Traumatismo dento alveolar	-
99	Não identificado	-



**MINISTÉRIO DA SAÚDE
SECRETARIA DE VIGILÂNCIA EM SAÚDE
DEPARTAMENTO DE VIGILÂNCIA EPIDEMIOLÓGICA
CENTRO DE INFORMAÇÕES ESTRATÉGICAS EM VIGILÂNCIA EM SAÚDE
GT-SINAN**

**SISTEMA DE INFORMAÇÃO DE AGRAVOS DE NOTIFICAÇÃO
DICIONÁRIO DE DADOS – SINAN NET – VERSÃO 5.0/Patch 5.1**

CAMPO DE PREENCHIMENTO OBRIGATÓRIO é aquele cuja ausência de dado impossibilita a inclusão da notificação ou da investigação no Sinan.
CAMPO ESSENCIAL é aquele que, apesar de não ser obrigatório, registra dado necessário à investigação do caso ou ao cálculo de indicador epidemiológico ou operacional.

Violência Interpessoal/Autoprovoçada

Nome do Campo	Campo	Tipo	Categoria	Descrição	Características	DBF
Nº da Notificação	nu_notificacao	varchar2(7)		Número da Notificação	Campo Chave para identificação do registro no sistema	NU_NOTIFIC
1. Tipo de Notificação	tp_notificacao	varchar2(1)	1 – Negativa 2 – Individual 3 – Surto 4 – Agregado	Identifica o tipo da notificação	Campo Obrigatório	TP_NOT
2. Agravo	co_cid	varchar2(4)	Tabela de agravos do sistema com códigos (classificação internacional de doenças –CID 10) e nomes dos agravos classificados como notificação compulsória (nacional, estadual ou municipal) e as síndromes (febre hemorragia aguda;respiratória aguda; diarreia aguda; sanguinolenta; neurológica aguda; insuficiência renal aguda; outras síndromes	Nome e código do agravo notificado segundo CID-10 (Centro Colaborador da OMS para a Classificação de Doenças em Português)	Campo Chave Preenchendo o código, a descrição é preenchida automaticamente, e vice-versa; Ao exportar, é retirado o ponto	ID_AGRAVO

3. Data da Notificação	dt_notificacao	date	dd/mm/aaaa	Data de preenchimento da ficha de notificação.	Campo Chave	DT_NOTIFIC
Semana epidemiológica da notificação (campo interno)	ds_semana_notificacao	varchar2(6)	Semanas do calendário epidemiológico padronizado.	Semana epidemiológica que o caso foi notificado.	Preenchida automaticamente, a partir da data de notificação (AAAASS)	SEM_NOT
(campo interno)	varchar(4)	varchar(4)		Ano da notificação	Variável interna preenchida pelo sistema a partir da data de notificação	NU_ANO
4. UF de Notificação	co_uf_notificacao	varchar2(2)	Tabela com Códigos e siglas padronizados pelo IBGE	Sigla da Unidade Federativa onde está localizada a unidade de saúde (ou outra fonte notificadora) que realizou a notificação. A sigla é uma variável que está associada ao código na tabela.	Campo Obrigatório	SG_UF_NOT
5. Município de Notificação	co_municipio_notificacao	varchar2(6)	Tabela com Código e nome dos municípios do cadastro do IBGE (tabela municipi.dbf)	Código do município onde está localizada a unidade de saúde (ou outra fonte notificadora) que realizou a notificação. O nome está associado ao código na tabela de municípios.	Campo Chave	ID_MUNICIP
(campo interno)	varchar2(4)	varchar2(4)	Campo com código da regional de saúde da tabela de município do sistema	Regional de saúde onde está localizado o município da unidade de saúde ou outra fonte notificadora	Sistema relaciona o campo município de notificação com o campo regional da tabela de municípios	ID_REGIONA
6. Unidade Notificadora	tp_unidade_notificacao	varchar2(1)	1-Unidade de Saúde 2- Unidade de Assistência Social 3-Estabelecimento de Ensino 4- Conselho Tutelar 5-Unidade de Saúde Indígena 6-Centro Especializado de Atendimento à Mulher 7-Outros	Setor de atuação da unidade notificadora	Campo Obrigatório Se for igual (=) 1. Unidade de Saúde ou 7. Outros, pular para o campo 8. Unidade de Saúde e preencher o campo 7. Nome da Unidade Notificadora automaticamente com a informação digitada no campo 8. Unidade de Saúde.	TP_UNI_EXT

7. Nome da Unidade Notificadora 7.1 Código	no_unidade_not	varchar2(15,0)	Código e nome das tabelas dos cadastros da Saúde (CNES), MDS, MEC, SDH, Saúde Indígena e SPM	Nome completo e código da unidade de que realizou a notificação do caso	Campo Obrigatório se campo 6. Unidade Notificadora for preenchido com 2. Unidade de Assistência Social, 3. Estabelecimento de Ensino, 4. Conselho Tutelar, 5. Unidade de Saúde Indígena ou 6. Centro Especializado de Atendimento à Mulher. Ao preencher o código, a descrição é preenchida automaticamente e vice-versa.	NM_UNI_EXT CO_UNI_EXT
8. Unidade de Saúde 8.1 Código CNES	co_unidade_notificacao	number (7,0)	Código e nome da tabela do cadastro Nacional de Estabelecimento de Saúde (CNES)	Nome completo e código da unidade de saúde da área de abrangência da unidade notificadora	Campo Obrigatório Ao preencher o código, a descrição é preenchida automaticamente e vice-versa.	ID_UNIDADE CNES_NOT
9. Data da ocorrência da violência	dt_ocor_viol	date	dd/mm/aaaa	Data da ocorrência da violência	Campo Obrigatório Data menor ou igual (<=) a Data de Notificação	DT_OCOR
Semana epidemiológica dos primeiros sintomas/diagnósticos. (campo interno)	ds_semana_sintoma	varchar2(6)	Semanas do calendário epidemiológico padronizado.	Semana epidemiológica dos números sintomas.	Preenchida automaticamente, a partir da data de primeiros sintomas data do diagnostico. (AAAASS)	SEM_PRI
10. Nome do Paciente	no_nome_paciente	varchar2(70)		Nome completo do paciente (sem abreviações)	Campo Obrigatório	NM_PACIENT

11. Data de nascimento	dt_nascimento	date	dd/mm/aaaa	Data de nascimento do paciente	<p>Campo Obrigatório Caso a idade não seja preenchida</p> <p>Uma vez preenchido é, calculada automaticamente a idade.</p>	DT_NASC
12. Idade	nu_idade	number(4)	<p>A composição da variável obedece o seguinte critério: 1º dígito:</p> <ol style="list-style-type: none"> 1. Hora 2. Dia 3. Mês 4. Ano <p>Ex: 3009 – nove meses, 4018 – dezoito anos</p>	quando não há data de nascimento a idade deve ser digitada segundo informação fornecida pelo paciente como aquela referida por ocasião da data da ocorrência ou na falta desse dado é registrada a idade aparente.	<p>Campo Obrigatório Caso a data de nascimento não esteja preenchida</p> <p>Preenchido automaticamente se Data de nascimento for preenchida</p> <p>Calculada entre Data de nascimento e Data de Ocorrência ;</p> <p>Campo composto pela unidade de medida de tempo e numero da idade.</p> <ul style="list-style-type: none"> - Se campo < 4007 (7 anos), campo escolaridade é preenchido com 10- Não se aplica. 	NU_IDADE_N
13. Sexo	tpsexo	varchar2(1)	<p>M- Masculino F- Feminino I- Ignorado</p>	Sexo do paciente	<p>Campo Obrigatório</p> <p>Se Sexo = Masculino, desabilitar o campo Gestante;</p>	CS_SEXO
14. Gestante	tpgestante	varchar2(1)	<p>Gestante</p> <ol style="list-style-type: none"> 1. 1º Trimestre 2. 2º Trimestre 3. 3º Trimestre 4. Idade gestacional ignorada 5. Não 6. Não se aplica 9. Ignorado 	Idade gestacional da paciente.	<p>Campo Obrigatório se sexo = F Se Sexo=Masculino ou idade menor ou igual a 10 anos, o campo é preenchido automaticamente com 6 Não se Aplica;</p>	CS_GESTANT

15. Raça/Cor	tp_raca_cor	varchar2(1)	1- branca 2- preta 3- amarela 4- parda 5- indígena 9 Ignorado	Considera-se cor ou raça declarada pela pessoa. 1. branca 2. preta 3. amarela (pessoa que se declarou de raça amarela) 4. parda (pessoa que se declarou mulata, cabocla, cafuza, mameluca ou mestiça de preto com pessoa de outra cor ou raça) 5. indígena (pessoa que se declarou indígena ou índia)	Campo Essencial	CS_RACA
16. Escolaridade	tp_escolaridade	varchar2(2)	43 Analfabeto 1. 1ª a 4ª série incompleta do EF 2. 4ª série completa do EF (antigo 1º grau) 3. 5ª à 8ª série incompleta do EF (antigo ginásio ou 1º grau) 4. Ensino fundamental completo (antigo ginásio ou 1º grau) 5. Ensino médio incompleto (antigo colegial ou 2º grau) 6. Ensino médio completo (antigo colegial ou 2º grau) 7. Educação superior incompleta 8. Educação superior completa 9. Ignorado 10. Não se aplica	Série e grau que a pessoa está freqüentando ou freqüentou considerando a última série concluída com aprovação ou grau de instrução do paciente por ocasião da notificação.	Campo Essencial Preenchido automaticamente com a Categoria 10-não se aplica quando idade menor a 7 anos de idade. - Quando caso notificado > 7 anos, campo não pode ser preenchido com categoria 10-não se aplica. .	CS_ESCOL_N
17. Nº. Cartão SUS	nu_cartao_sus	varchar2(15)		Número do cartão do Sistema Único de Saúde (SUS) do paciente		ID_CNS_SUS
18. Nome da mãe	no_nome_mae	varchar2(60)		Nome completo da mãe do paciente (sem abreviações)	Campo Essencial	NM_MAE_PAC
19. UF	co_uf_residencia	varchar2(2)	Tabela com Códigos e siglas	Sigla da Unidade Federada de	Campo Obrigatório se	SG_UF

			padronizados pelo IBGE (tabela municipi.dbf)	residência do paciente por ocasião da notificação. A sigla é uma variável que está associada ao código na tabela.	residente no Brasil. Ao digitar sigla da UF, o campo País é preenchido automaticamente com “Brasil”. Se nenhuma UF for selecionada, o sistema pula automaticamente para seleção de outro país que não o Brasil.	
20. Município de residência	co_municipio_residencia	varchar2(6)	Tabela com Códigos e nomes padronizados pelo IBGE	Código do município de residência do caso notificado. O nome está associado ao código na tabela de municípios.	Campo Obrigatório quando UF é digitada. São exibidos somente os Municípios pertencentes à UF selecionada no campo anterior. Permite digitação do nome do município ou do código IBGE. Quando digitado o nome, o código é preenchido automaticamente e vice-versa.	ID_MN_RESI
(campo interno)		varchar2(4)	Campo com código da regional de saúde da tabela de município do sistema	Regional de saúde onde está localizado o município de residência do paciente por ocasião da notificação	Sistema relaciona o campo município de notificação com o campo regional da tabela de municípios	ID_RG_RESI
21. Distrito	co_distrito_residencia	varchar(9)	Códigos e nomes padronizados segundo Tabela disponibilizada pelo sistema para cadastramento pelo gestor municipal do Sinan	Nome e respectivo código do distrito de residência do paciente por ocasião da notificação.	Campo Essencial Filtra apenas os distritos pertencentes ao Município selecionado no campo anterior.	ID_DISTRI
22. Bairro	co_bairro_residen	number(8)	Códigos seqüenciais da tabela	Código do bairro de residência	Campo Essencial	ID_BAIRRO

	cia		de localidade	do paciente por ocasião da notificação. Serão exibidos apenas os Bairros pertencentes ao Município selecionado no campo anterior.	Se existir bairro cadastrado será gravado o código do bairro no banco.	
22. Nome do Bairro	no_bairro_residencia	varchar2(60)		Nome do bairro de residência	Se o bairro não estiver cadastrado na tabela municipal do Sinan será gravado o nome digitado no campo Bairro e não será gravado nenhum código.	NM_BAIRRO
23. Logradouro (rua, avenida...)	co_logradouro_residencia	number(8)	Tipo e nome do logradouro digitado na entrada de dados ou, se disponível selecionado em tabela de logradouros do município (código e descrição)	Identificação do tipo (avenida, rua, travessa, etc.) título e nome do logradouro Dados do endereço de residência do paciente por ocasião da notificação	Campo Essencial Se o logradouro for de município diferente do município de residência, será gravado o nome digitado no campo e não será gravado nenhum código.	ID_LOGRADO
	varchar2(60)	varchar2(60)	Descrição digitada quando não há tabela no sistema.	Identificação do tipo (avenida, rua, travessa, etc.) título e nome do logradouro Dados do endereço de residência do paciente por ocasião da notificação (av., rua, etc.).	Campo Essencial Esse campo será digitado quando não houver Logradouro	NM_LOGRADO
24. Número do logradouro	nu_residencia	varchar2(6)		Nº. do logradouro (nº. da casa ou do edifício) Dados do endereço de residência do paciente por ocasião da notificação	Campo Essencial	NU_NUMERO
25. Complemento do logradouro	ds_complemento_residencia	varchar2(30)		Complemento do logradouro (ex. Bloco D, apto. casa, etc.) Dados do endereço de residência do paciente por ocasião da notificação	Campo Essencial	DS_COMPL

26. Geo campo 1	co_geo_campo_1	number(8)	Selecionado em tabela de localidade (código e descrição) Tabela com Códigos e nomes padronizados segundo Tabela disponibilizada pelo sistema para importação pelo gestor municipal do Sinan de tabela georeferenciada	Campo para Geo-Referenciamento		ID_GEO1
27. Geo campo 2	co_geo_campo_2	number(8)	Selecionado em tabela de localidade (código e descrição) Tabela com Códigos e nomes padronizados segundo Tabela disponibilizada pelo sistema para importação pelo gestor municipal do Sinan de tabela georeferenciada	Campo para Geo-Referenciamento		ID_GEO2
28. Ponto de referencia	ds_referencia_residencia	varchar2(70)		Ponto de referência para facilitar a localização da residência do paciente	Campo Essencial	DS_REF_RES
29. CEP de residência	nu_cep_residencia	varchar2(8)		CEP de residência do paciente.	Campo Essencial	NU_CEP
30. (DDD)	nu_ddd_residencia	varchar2(2)		Código de residência do paciente	Campo Essencial	DDD
30. Telefone	nu_telefone_residencia	varchar2(9)		Telefone de residência do paciente	Campo Essencial	FONE
31. Zona	tp_zona_residencia	varchar2(1)	1. urbana 2. rural 3. periurbana 9. ignorado	Zona de residência do paciente por ocasião da notificação	Campo Essencial	ZONA
32. País (se residente fora do Brasil)	co_pais_residencia	varchar(4)	Tabela com código e descrição de países.	País onde residia o paciente por ocasião da notificação	Campo Obrigatório Se UF for digitada este campo é preenchido automaticamente com "Brasil".	ID_PAIS
Não Listar/ Não Contar	tp_duplicidade	varchar2(1)	0 ou branco – Não identificado 1. Não é duplicidade (não listar) 2. Duplicidade (não contar)	Identifica duplicidade	Categorias de 1 e 2 atribuídas pelo usuário do sistema na rotina de	NDUPLIC_N

					<p>duplicidade para identificar duplicidades que não devem ser listadas no relatório de duplicidade ou não devem ser computadas na incidência do agravo.</p> <p>Quando o registro estava marcado para <i>não listar</i> e retornou para a rotina de duplicidade, é atribuída a categoria 0.</p>	
Vinculação	st_vincula	varchar2(1)	0 ou branco – Não Vinculado 1 – Vinculado	Indica se a notificação foi vinculada	Categoria 1 atribuída pelo sistema após vinculação de notificações de hanseníase ou tuberculose.	IN_VINCULA
33. Nome social						
34. Ocupação	co_cbo_ocupacao	Varchar2(6)	Categorias da tabela (Indicar aqui qual a tabela que é utilizado no Sinan NET)	Informar a atividade exercida pelo paciente no setor formal, informal ou autônomo ou sua última atividade exercida quando paciente for desempregado. O ramo de atividade econômica do paciente refere-se as atividades econômicas desenvolvidas nos processos de produção do setor primário (agricultura e extrativismo); secundário (indústria) ou terciário (serviços e comércio).		ID_OCUPA_N
35. Situação conjugal	tp_situacao_conjugal	varchar2(1)	<ol style="list-style-type: none"> 1. Solteiro 2. Casado/ União consensual 3. Viúvo 4. Separado 8. Não se aplica 	Situação conjugal do paciente	Campo Essencial Categoria 8 se idade do paciente for <=9 anos	SIT_CONJUG

			9. Ignorado			
36. Orientação sexual	tp_orientacao_sexual	varchar2(1)	<ol style="list-style-type: none"> 1. Heterossexual 2. Homossexual (gay/lésbica) 3. Bissexual 8. Não se aplica 9. Ignorado 	Orientação sexual do paciente	Campo Obrigatório se >=10 anos Categoria (8) Não se Aplica se idade do paciente for <=9 anos	ORIENT_SEX
37. Identidade de Gênero	identidade_genero	varchar2(1)	<ol style="list-style-type: none"> 1. Travesti 2. Transexual Mulher 3. Transexual Homem 8. Não se aplica 9. Ignorado 	Identidade de gênero do paciente	Campo Obrigatório se >=10 anos Categoria (8) Não se Aplica se idade do paciente for <=9 anos Se o campo 13 Sexo for igual (=) a Feminino (F) não permitir a categoria (4) Travesti	IDENT_GEN
38. Possui algum tipo de deficiência/ transtorno	tp_deficiencia_transtorno	varchar2(1)	<ol style="list-style-type: none"> 1. Sim 2. Não 9. Ignorado 	Se o paciente possui algum tipo de deficiência/ transtorno	Campo Essencial Se 2 ou 9, pular para o campo 40 (UF de ocorrência)	DEF_TRANS
39. Deficiência física	tp_deficiencia_fisica	varchar2(1)	<ol style="list-style-type: none"> 1. Sim 2. Não 8. Não se aplica 9. Ignorado 	Se o paciente possui algum tipo de deficiência física	Campo Essencial Se campo 38 for = 2 ou 9 preencher automaticamente com categoria = 8 (não se aplica). Se campo 38 = 1, não aceitar categoria = 8 (não se aplica).	DEF_FISICA
39. Deficiência intelectual	tp_deficiencia_intelectual	varchar2(1)	<ol style="list-style-type: none"> 1. Sim 2. Não 8. Não se aplica 9. Ignorado 	Se o paciente possui algum tipo de deficiência mental	Campo Essencial Se campo 38 for = 2 ou 9 preencher	DEF_MENTAL

					automaticamente com categoria = 8 (não se aplica). Se campo 38 = 1, não aceitar categoria = 8 (não se aplica).	
39. Deficiência visual	tp_deficiencia_visual	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Se o paciente possui algum tipo de deficiência visual	Campo Essencial Se campo 38 for = 2 ou 9 preencher automaticamente com categoria = 8 (não se aplica). Se campo 38 = 1, não aceitar categoria = 8 (não se aplica).	DEF_VISUAL
39. Deficiência auditiva	tp_deficiencia_auditiva	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Se o paciente possui algum tipo de deficiência auditiva	Campo Essencial Se campo 38 for = 2 ou 9 preencher automaticamente com categoria = 8 (não se aplica). Se campo 38 = 1, não aceitar categoria = 8 (não se aplica).	DEF_AUDITI
39. Transtorno mental	tp_transtorno_mental	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Se o paciente possui algum tipo de transtorno mental	Campo Essencial Se campo 38 for = 2 ou 9 preencher automaticamente com categoria = 8 (não se aplica). Se campo 38 = 1, não aceitar categoria = 8 (não se aplica).	TRAN_MENT
39. Transtorno de	tp_transtorno_cognitivo	varchar2(1)	1. Sim	Se o paciente possui algum tipo	Campo Essencial	TRAN_COMP

comportamento	mportamento		2. Não 8. Não se aplica 9. Ignorado	de transtorno de comportamento	Se campo 38 for = 2 ou 9 preencher automaticamente com categoria = 8 (não se aplica). Se campo 38 = 1, não aceitar categoria = 8 (não se aplica).	
39. Outras deficiências/síndromes	tp_deficiencia_sindrome	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Se o paciente possui algum tipo de outras deficiências/síndromes	Campo Essencial Se campo 38 for = 2 ou 9 preencher automaticamente com categoria = 8 (não se aplica). Se campo 38 = 1, não aceitar categoria = 8 (não se aplica).	DEF_OUT
39. Outras deficiências/síndromes (especificar)	ds_sindrome	varchar2(30)		Especificar qual o tipo de outras deficiências/síndromes	Campo obrigatório se campo 39 Outras deficiências/síndromes for = 1	DEF_ESPEC
40. UF de ocorrência	co_uf_ocorrencia	varchar2(2)	Tabela com Código dos estados padronizados pelo IBGE	Sigla da Unidade Federada de ocorrência do evento por ocasião da notificação	Campo obrigatório	SG_UF_OCOR
41. Município de ocorrência	co_municipio_residencia	varchar2(6)	Tabela com Códigos e nomes padronizados pelo IBGE	Código e nome do município de ocorrência do evento notificado.	Campo obrigatório São exibidos somente os municípios pertencentes à UF selecionada no campo anterior. Digitação do nome do município ou do código: quando digitado o nome, o código é preenchido automaticamente e vice-versa.	ID_MN_OCOR

42. Distrito de ocorrência	co_distrito_ocorrencia	number(8)	Tabela com Códigos e nomes padronizados segundo Tabela disponibilizada pelo sistema para cadastramento pelo gestor municipal do Sinan Ou tabela nacional carregada no sistema	Nome e respectivo código do distrito de ocorrência do evento	Filtra apenas os distritos pertencentes ao Município selecionado no campo 41;	ID_DIS_OCOR
43. Bairro de ocorrência	co_bairro_ocor	number(8)	Tabela Códigos e nomes padronizados segundo Tabela disponibilizada pelo sistema para cadastramento pelo gestor municipal do Sinan	Nome e respectivo código do bairro de ocorrência do evento. Serão exibidos apenas os Bairros pertencentes ao Município selecionado no campo anterior.	Campo essencial Se o bairro for de município diferente do município de ocorrência, será gravado o nome digitado no campo bairro (no_bairro_ocor) e não será gravado nenhum código.	ID_BA_OCOR
43.Nome do Bairro	no_bairro_ocor	varchar2(60)			Esse campo será digitado quando não houver bairro cadastrado	NM_BA_OCOR
44. Logradouro (rua, avenida...) de ocorrência	co_logradouro_ocorrencia	number(8)	Tipo e nome do logradouro digitado na entrada de dados ou, se disponível selecionado em tabela de logradouros do município (código e descrição)	Identificação do tipo (avenida, rua, travessa, etc.) título e nome do logradouro Dados do endereço de ocorrência do evento	Se o logradouro for de município diferente do município de residência, será gravado o nome digitado no campo e não será gravado nenhum código.	ID_LOG_OCO
44. Logradouro	no_logradouro_ocorrencia	varchar2(50)	Descrição digitada quando não há tabela no sistema.	Identificação do tipo (avenida, rua, travessa, etc.) Título e nome do logradouro Dados do endereço de ocorrência do evento (av., rua, etc)		NO_LOG_OCO
45. Número do logradouro de ocorrência	nu_ocorrencia	varchar2(6)		Nº. do logradouro (nº. da casa ou do edifício) Dados do endereço de ocorrência do evento		NM_LOG_RES
46. Complemento do logradouro de ocorrência	ds_complemento_ocorrencia	varchar2(30)		Complemento do logradouro (ex. Bloco D, apto. casa, etc.) Dados do endereço de		DS_COMP_OCOR

				ocorrência do evento		
47. Geo campo 3	co_geo_campo_3	number(8)	Selecionado em tabela de localidade (código e descrição) Tabela com Códigos e nomes padronizados segundo Tabela disponibilizada pelo sistema para cadastramento pelo gestor municipal do Sinan	Campo para Geo-Referenciamento		ID_GEO3
48. Geo campo 4	no_geo_campo_4	number(8)	Selecionado em tabela de localidade (código e descrição) Tabela com Códigos e nomes padronizados segundo Tabela disponibilizada pelo sistema para cadastramento pelo gestor municipal do Sinan	Campo para Geo-Referenciamento		ID_GEO4
49. Ponto de referencia da ocorrencia	ds_referencia_oco rrencia	varchar2(70)		Ponto de referência para facilitar a localização do local de ocorrência do evento		DS_REF_OCO R
50. Zona de ocorrencia	tp_zona_ocorrenc ia	varchar2(1)	1. Urbana 2. Rural 3. Periurbano 9. Ignorado	Zona de ocorrencia do evento por ocasião da notificação		ZONA_OCOR
51. Hora de ocorrência	hr_ocorrencia	hora (hh:mm)		Informar a hora em que a violência ocorreu considerando o intervalo de 00:00 a 23:59s	Campo essencial	HORA_OCOR
52. Local de ocorrência	tp_local_ocorrenc ia	varchar2(2)	01. Residência 02. Habitação coletiva 03. Escola 04. Local de prática esportiva 05. Bar ou similar 06. Via publica 07. Comércio/Serviços 08. Industrias/ construção 09. Outro 99. Ignorado	Informar o local de ocorrência do evento notificado	Campo obrigatório Se categoria selecionada for diferente de 09, pular para questão 53	LOCAL_OCOR
52. Local de ocorrência - Outro (especificar)	Ds_outro_tp_loca l_ocorrencia	varchar2(30)		Especificar outro local de ocorrência	Campo obrigatório se campo 52 Local de Ocorrência for igual a 9-Outro	LOCAL_ESPE

53. Ocorreu outras vezes?	st_ocorreu_outra_vez	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se a violência é de repetição	Campo Essencial	OUT_VEZES
54. A lesão foi autoprovocada?	st_lesao_autoprovocada	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se a lesão foi autoprovocada	Campo essencial	LES_AUTOP
55. Violência motivada por:	tp_violencia_motivada	varchar2(1)	01. Sexismo 02. Homofobia/Lesbofobia Bifobia/Transfobia 03. Racismo 04. Intolerância religiosa 05. Xenofobia 06. Conflito geracional 07. Situação de rua 08. Deficiência 09. Outros 88. Não se aplica 99. Ignorado	Informar se violência tem relação com características da vítima	Campo obrigatório	VIOL_MOTIV
56. Tipo de Violência - Física	st_violencia_fisica	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu violência física	Campo obrigatório	VIOL_FISIC
56. Tipo de Violência - psicológica/ moral	st_violencia_psico_moral	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu violência psicológica/moral	Campo obrigatório	VIOL_PSICO
56. Tipo de Violência - Tortura	st_tortura	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu tortura	Campo obrigatório	VIOL_TORT
56. Tipo de Violência - Sexual	st_violencia_sexual	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu violência sexual	Campo obrigatório	VIOL_SEXU
56. Tipo de Violência - Tráfico de seres humanos	st_trafico_ser_humano	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu tráfico de seres humanos	Campo obrigatório	VIOL_TRAF
56. Tipo de Violência -	st_violencia_finan	varchar2(1)	1. Sim	Informar se ocorreu violência	Campo obrigatório	VIOL_FINAN

financeira/econômico	_economica		2. Não 9. Ignorado	financeira/econômica		
56. Tipo de Violência - Negligencia/ abandono	st_neglicencia_abandono	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu negligencia/abandono	Campo obrigatório	VIOL_NEGLI
56. Tipo de Violência - Trabalho infantil	st_trabalho_infantil	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu trabalho infantil	Campo obrigatório	VIOL_INFAN
56. Tipo de Violência - Intervenção legal	st_intervencao_legal	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu intervenção legal	Campo obrigatório	VIOL_LEGAL
56. Tipo de Violência - Outros	st_outro_tipo_violencia	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se ocorreu outro tipo de violência	Campo obrigatório Se categoria=2 ou 9, pular para o campo 57. Meio de agressão Força Corporal/Espancamento	VIOL_OUTR
56. Outros tipos de violência (especificar)	ds_outro_tipo_violencia	varchar2(30)		Especificar que outro tipo de violência ocorreu.	Campo Obrigatório se campo 56-Outros for = 1.	VIOL_ESPEC
57. Meio de Agressão Força corporal/ espancamento	st_forca_corporal_espancamento	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de força corporal	Campo Essencial	AG_FORCA
57. Meio de Agressão Enforcamento	st_enforcamento	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de enforcamento	Campo Essencial	AG_ENFOR
57. Meio de Agressão Objeto contundente	st_objeto_contudente	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de objeto contundent	Campo Essencial	AG_OBJETO
57. Meio de Agressão Objeto perfuro-cortante	st_objeto_perfuro_cortante	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de obj. perfuro cortant	Campo Essencial	AG_CORTE
57. Meio de Agressão Substancia/objeto	st_subst_objeto_quente	varchar2(1)	1. Sim 2. Não	Informar se o meio de agressão foi através de subst/obj quente	Campo Essencial	AG_QUENTE

quente			9. Ignorado			
57. Meio de Agressão Envenenamento, Intoxicação	st_envenenamento	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de envenenamento	Campo Essencial	AG_ENVEN
57. Meio de Agressão Arma de fogo	st_arma_fogo	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de arma de fogo	Campo Essencial	AG_FOGO
57. Meio de Agressão Ameaça	st_ameaca	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de ameaça	Campo Essencial	AG_AMEACA
57. Meio de Agressão Outro	st_outro_meio_agressao	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o meio de agressão foi através de outro	Se categoria selecionada for =2 ou 9, pular para o campo 58. Se ocorreu violência sexual, qual tipo? - Assédio sexual	AG_OUTROS
57. Meio de Agressão Outro meio de agressão (especificar)	ds_outro_meio_agressao	varchar2(30)		Especificar qual outro meio de agressão sofreu.	Campo Obrigatório se campo 57-Outro for = 1.	AG_ESPEC
58. Se ocorreu violência sexual, qual tipo? - Assédio sexual	st_ocorreu_violencia_assedio	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu assédio sexual	Campo essencial Habilitado se campo 56- violência sexual for=1. Categoria 8 (não se aplica) se campo 56 – tipo de violência sexual for = 2 ou 9	SEX_ASSEDI
58. Se ocorreu violência sexual, qual tipo? - Estupro	st_ocorreu_violencia_estupro	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu estupro	Campo essencial Habilitado se campo 56- violência sexual for=1. Categoria 8 (não se aplica) se campo 56 – tipo de violência sexual for = 2 ou	SEX_ESTUPR

					9	
58. Se ocorreu violência sexual, qual tipo? - Pornografia infantil	st_ocorreu_violencia_porno_inf	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu Pornografia infantil	Campo essencial Habilitado se campo 56- violência sexual for=1. Categoria 8 (não se aplica) se campo 56– tipo de violência sexual for = 2 ou 9	SEX_PORNO
58. Se ocorreu violência sexual, qual tipo? - Exploração sexual	st_ocorreu_violencia_exp_sexual	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu exploração sexual	Campo essencial Habilitado se campo 56- violência sexual for=1. Categoria 8 (não se aplica) se campo 56 – tipo de violência sexual for = 2 ou 9	SEX_EXPLO
58. Se ocorreu violência sexual, qual tipo? - Outro	st_ocorreu_violencia_sexual_outro	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu outro tipo de violência sexual	Campo essencial Habilitado se campo 56- violência sexual for=1. Categoria 8 (não se aplica) se campo 56 – tipo de violência sexual for = 2 ou 9 Se categoria selecionada for =2, 8 ou 9, pular para	SEX_OUTRO

					campo 59. Procedimento realizado.	
58. Se ocorreu violência sexual, qual tipo? - Outro tipo de violência sexual (especificar)	ds_ocorreu_violencia_sexual_outro	varchar2(30)		Especificar qual outro tipo de violência sexual	Campo Obrigatório se campo 58. Se ocorreu violência sexual, qual tipo? - Outro =1	SEX_ESPEC
59. Procedimento realizado – Profilaxia DST	st_proc_profilaxia_dst	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu profilaxia DST	Campo essencial Habilitado se campo 56- violência sexual for=1. Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente	PROC_DST
59. Procedimento realizado – Profilaxia HIV	st_proc_profilaxia_hiv	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu profilaxia HIV	Campo essencial Habilitado se campo 56- violência sexual for=1. Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente	PROC_HIV
59. Procedimento realizado – Profilaxia Hepatite B	st_proc_profilaxia_hepatite_b	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu profilaxia hepatite B	Campo essencial Habilitado se campo 56- violência sexual for=1. Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente	PROC_HEPB
59. Procedimento realizado – Coleta de sangue	st_proc_coleta_sangue	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu coleta de sangue	Campo essencial Habilitado se campo 56- violência sexual for=1. Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente	PROC_SANG
59. Procedimento realizado – Coleta de sêmen	st_proc_coleta_semen	varchar2(1)	1. Sim 2. Não 8. Não se aplica	Informar se ocorreu coleta de semen	Campo essencial Habilitado se campo 56- violência sexual for=1.	PROC_SEMEN

			9. Ignorado		Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente	
59. Procedimento realizado – Coleta de secreção vaginal	st_proc_coleta_secrecao_vagina	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu coleta de secreção vaginal	Campo essencial Habilitado se campo 56- violência sexual for=1. Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente Se sexo igual a Masculino (M) preencher com não se aplica (8)	PROC_VAGINA
59. Procedimento realizado – Contraceção de emergência	st_proc_contracepcao_emergenc	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu contracepção de emergência	Campo essencial Habilitado se campo 56- violência sexual for=1. Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente Se sexo igual a Masculino (M) preencher com não se aplica (8)	PROC_CONTR
59. Procedimento realizado - Aborto previsto em lei	st_porc_aborto_previsto_lei	varchar2(1)	1. Sim 2. Não 8. Não se aplica 9. Ignorado	Informar se ocorreu aborto previsto em lei	Campo essencial Habilitado se campo 56- violência sexual for=1. Se campo 56 violência sexual for = 2 ou 9 preencher com categoria 8 automaticamente Se sexo igual a Masculino	PROC_ABORT

					(M) preencher com não se aplica (8)	
60. Numero de envolvidos	tp_numero_envolvido	varchar2(1)	1. Um 2. Dois ou mais 9. Ignorado	Informar o número de envolvidos na violência		NUM_ENVOLV
61. Relação com a pessoa atendida – Pai	st_provavel_autor_pai	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o pai	Campo obrigatório	REL_PAI
61. Relação com a pessoa atendida – Mãe	st_provavel_autor_mae	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era a mãe	Campo obrigatório	REL_MAE
61. Relação com a pessoa atendida – Padrasto	st_provavel_autor_padrasto	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o padrasto	Campo obrigatório	REL_PAD
61. Relação com a pessoa atendida – Madrasta	st_provavel_autor_madrasta	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era a madrasta	Campo obrigatório	REL_MAD
61. Relação com a pessoa atendida – Cônjuge	st_provavel_autor_conjuge	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o cônjuge	Campo obrigatório	REL_CONJ
61. Relação com a pessoa atendida – Ex-cônjuge	st_provavel_autor_ex_conjuge	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o ex-cônjuge	Campo obrigatório	REL_EXCON
61. Relação com a pessoa atendida – Namorado(a)	st_provavel_autor_namorado	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o namorado(a)	Campo obrigatório	REL_NAMO
61. Relação com a pessoa atendida – Ex-namorado(a)	st_provavel_autor_ex_namorado	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o ex-namorado(a)	Campo obrigatório	REL_EXNAM
61. Relação com a pessoa atendida – Filho(a)	st_provavel_autor_filho	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o (a) provável autor (a) da agressão eram filho ou filha	Campo obrigatório	REL_FILHO

61. Relação com a pessoa atendida – Irmão (ã)	st_provavel_autor_irmao	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era irmão ou irmã	Campo obrigatório	REL_IRMÃO
61. Relação com a pessoa atendida – Amigos/conhecidos	st_provavel_autor_amigo_conhecido	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão eram amigo(s)/conhecido(s)	Campo obrigatório	REL_CONHEC
61. Relação com a pessoa atendida – Desconhecido	st_provavel_autor_desconhecido	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era desconhecido	Campo obrigatório	REL_DESCO
61. Relação com a pessoa atendida – cuidador	st_provavel_autor_cuidador	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o cuidador	Campo obrigatório	REL_CUIDA
61. Relação com a pessoa atendida – patrão/chefe	st_provavel_autor_patrao_chefe	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era o patrão/chefe	Campo obrigatório	REL_PATRAO
61. Relação com a pessoa atendida – pessoa com relação institucional	st_provavel_autor_institucional	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era pessoa com relação institucional	Campo obrigatório	REL_INST
61. Relação com a pessoa atendida – policial/agente da lei	st_provavel_autor_agente	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era policial/agente da lei	Campo obrigatório	REL_POL
61. Relação com a pessoa atendida – Própria pessoa	st_prov_autor_propria	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão era a própria pessoa	Campo obrigatório	REL_PROPRI
61. Relação com a pessoa atendida – Outros	st_provavel_autor_outro	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão tinha outro tipo de relação	Campo obrigatório Se categoria=2 ou 9, pular para campo 62. Sexo do provável autor da agressão	REL_OUTROS
61. Relação com a pessoa atendida – Outros (especificar)	ds_provavel_autor_outro	varchar2(30)		Especificar outro tipo de relação com a pessoa atendida	Campo Obrigatório se campo 61. Relação com a pessoa atendida – Outros =1	REL_ESPEC

62. Sexo do provável autor da violência	tp_sexo_provavel_autor	varchar2(1)	1. Masculino 2. Feminino 3. Ambos os sexos 9. Ignorado	Informar o sexo do provável autor da agressão	Campo obrigatório	AUTOR_SEXO
63. Suspeita de uso de álcool	st_suspeita_uso_alcool	varchar2(1)	1. Sim 2. Não 9. Ignorado	Informar se o provável autor da agressão tinha suspeita de uso de álcool		AUTOR_ALCO
64. Ciclo de vida do Principal provável autor da violência	Tp_ciclo_de_vida_autor	varchar2(1)	1. Criança 2. Adolescente 3. Jovem 4. Pessoa adulta 5. Pessoa idosa 9. Ignorado	Informar o ciclo de vida do provável autor da agressão	Campo obrigatório	CICL_VID_AUTOR
65. Encaminhamento - Rede da Saúde	st_saude	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento no setor da rede da saúde	Campo obrigatório	ENC_SAUDE
65. Encaminhamento - Rede da Assistência Social	st_assistencia_social	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento no setor da rede Assistência Social	Campo obrigatório	ASSIST_SOC
65. Encaminhamento - Rede de Educação	st_rede_de_educacao	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento no setor da rede Educação	Campo obrigatório	REDE_EDUCA
65. Encaminhamento - Rede de Atendimento à Mulher	st_rede_atendimento_mulher	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento no setor da rede atendimento à mulher	Campo obrigatório	ATEND_MULHER
65. Encaminhamento - Conselho Tutelar	st_conselho_tutelar	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para o Conselho Tutelar	Campo obrigatório	CONS_TUTEL
65. Encaminhamento - Conselho do Idoso	st_conselho_idoso	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para o Conselho do Idoso	Campo obrigatório	CONS_IDO
65. Encaminhamento - Delegacia de Atendimento ao Idoso	st_del_atendimento_idoso	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para delegacia de atendimento ao idoso	Campo obrigatório	DELEG_IDOSO
65. Encaminhamento - Centro de Referência dos Direitos Humanos	St_centro_ref_direitos_humanos	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para centro de referência dos direitos Humanos	Campo obrigatório	DIR_HUMAN
65. Encaminhamento -	st_mp		1-Sim	Informar se houve	Campo obrigatório	MPU

Ministério Público		varchar2(1)	2-Não 9-Ignorado	encaminhamento para ministério Público		
65. Encaminhamento - Delegacia Especializada de Proteção à Criança e Adolescente	st_del_proteção_criança_adolescente	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para Delegacia Especializada de Proteção à Criança e Adolescente	Campo obrigatório	DELEG_CRIA
65. Encaminhamento - Delegacia de Atendimento à Mulher	st_del_atendimento_mulher	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para Delegacia de atendimento à mulher	Campo obrigatório	DELEG_MULH
65. Encaminhamento - Outras delegacias	st_outra_delegacia	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para outras delegacias	Campo obrigatório	DELEG
65. Encaminhamento - Justiça da infância e da Juventude	st_justiça_infância_juventude	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para justiça da infância e da juventude	Campo obrigatório	INFAN_JUV
65. Encaminhamento - Defensoria pública	st_defensoria_publica	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se houve encaminhamento para defensoria pública	Campo obrigatório	DEFEN_PUBL
66. Violência relacionada ao trabalho	st_violencia_rel_trabalho	varchar2(1)	1-Sim 2-Não 9-Ignorado	Informar se ocorreu violência relacionada ao trabalho	Campo Essencial Se categoria=2 ou 9 pular para o campo 68. Circunstância da lesão	REL_TRAB
67. Se sim, foi emitida a comunicação de acidente de trabalho (CAT)	st_emitida_cat	varchar2(1)	1-Sim 2-Não 8-Não se aplica 9-Ignorado	Informar se foi emitida a CAT, caso a violência seja relacionada ao trabalho	Categoria=8 se campo 66. Violência relacionada ao trabalho for = 2 ou 9 Se campo 66. Violência relacionada ao trabalho for = 1 não permitir a categoria 8. Não se aplica	REL_CAT
68. Circunstância da lesão	co_cid_lesao	varchar2(5)	Tabela de agravos do sistema com códigos (classificação internacional de doenças - CID 10) e nomes dos agravos	Nome e código do agravo notificado segundo CID-10 CAPITULO XX (VO1 a Y98)	Campo Essencial	CIRC_LESAO
69. Data de encerramento	dt_encerramento	date	Dd/mm/aaaa	data de encerramento do caso.	Campo ≥ data da notificação	DT_ENCERRA
Observações adicionais	ds_observacao					DS_OBS

9. Anexos

Anexo I

Modelo da Declaração de Óbito


República Federativa do Brasil
Ministério da Saúde

Declaração de Óbito Nº _____

I - Identificação

Centro Código Registro Data
 Município UF Censo

II - Identificação do falecido

Tipo de Óbito Obito Hora RHC Nacionalidade
 Nome do falecido Nome do pai Nome da mãe

III - Identificação do local de ocorrência

Data de Nascimento Idade completa Menores de 1 ano Sexo Raça/cor Estado civil Escolaridade Ocupação habitual e ramo de atividade
 Logradouro (Rua, praça, avenida etc.) Código Número Complemento CEP

IV - Identificação do estabelecimento

Local de ocorrência do óbito Estabelecimento Endereço da ocorrência, se fora do estabelecimento ou da residência (Rua, praça, avenida etc.)
 Bairro/Distrito Código Município de ocorrência Código UF

V - Informações sobre a mãe e o parto

Mãe Escolaridade Ocupação habitual e ramo de atividade da mãe Número de filhos vivos (Nasceram vivos) Mortos
 Duração da gestação (Em semanas) Tipo de Gravidez Tipo de parto Morte em relação ao parto Piso ao nascer Número Declar. de Nascidos Vivos

VI - Diagnóstico e causas da morte

Causas da morte Diagnóstico confirmado por: Cirurgia? Necrópsia?
 Parte I: Causa da morte CID Parte II: Outras condições significativas que contribuíram para a morte

VII - Médico

Nome do médico CRM O médico que assinou atendeu ao falecido?
 Meio de contato (Telefone, fax, e-mail etc.) Data do atestado Assinatura

VIII - Circunstâncias da morte

Tipo de morte Acidente de trabalho Fonte da informação
 Descrição sumária do evento, incluindo o tipo de local de ocorrência

IX - Assinatura do Médico

Declarante Testemunhas

Versão 0098-01
 Via do Operário do Saúde (Brasão), Centro de Registro civil (Jardim)
 Unidade de Saúde (FUS)