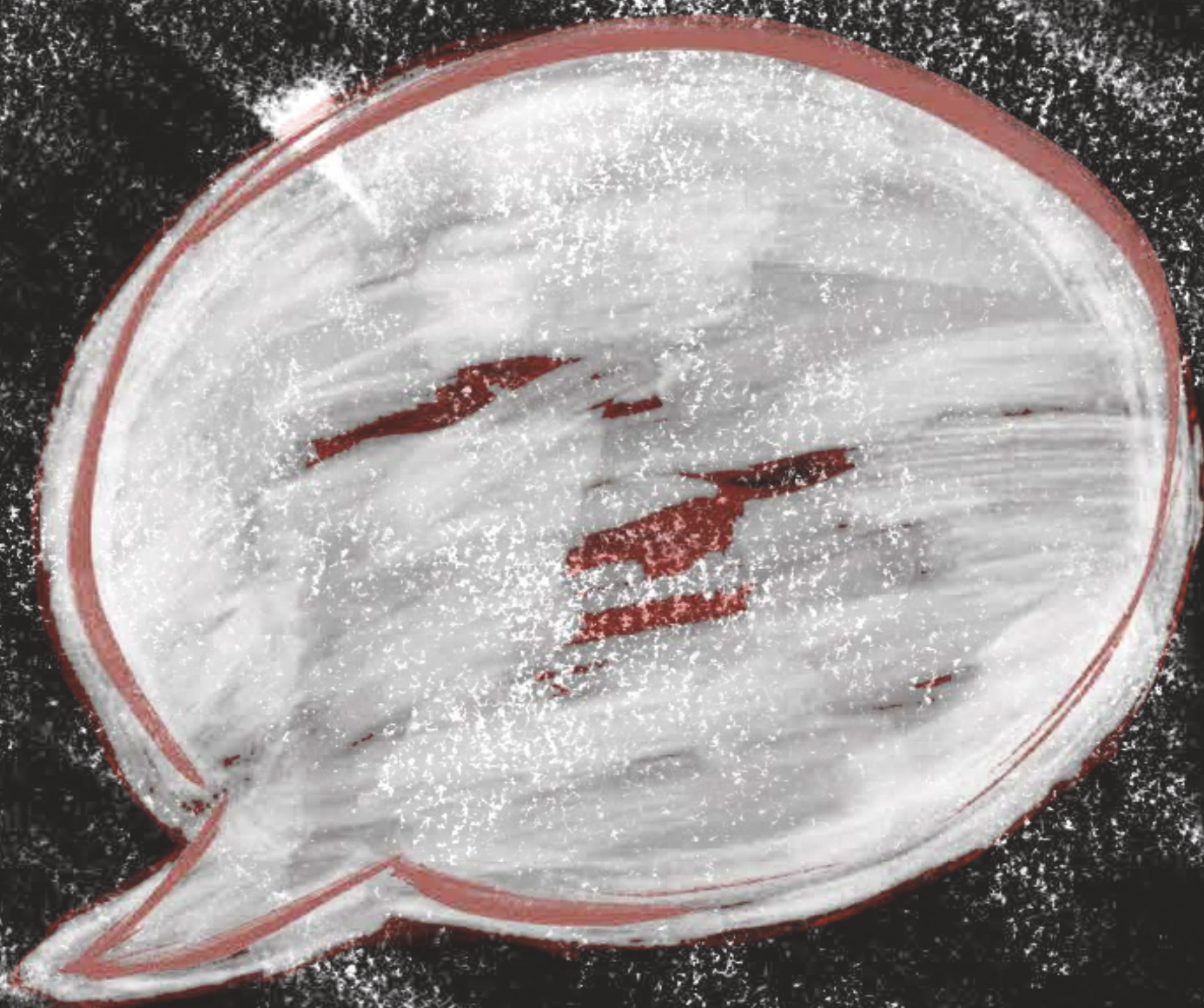


**KNOWING
ONLINE
HATE SPEECH**

KNOWLEDGE + AWARENESS = TACKLING HATE



**DESMASCARAR
O DISCURSO DE ÓDIO:
O CONHECIMENTO É A NOSSA ARMA**



ÍNDICE

- Sobre o que é esta brochura? **3**
- O que é discurso de ódio online? **4**
- Sabiam que existem diferentes tipos de discurso de ódio online? **6**
- Que tipo de discurso de ódio online é mais comum em Portugal? **8**
- Como podem identificar o discurso de ódio online? **9**
- O que é o contra-discurso online? **15**
- Como podem combater o discurso de ódio? **19**
- Comunidades racializadas - SOS Racismo **21**
- Comunidades LGBTI+ - ILGA **22**
- Comunidades migrantes - Casa do Brasil em Lisboa **23**
- Testem o protótipo! **24**
- Referências **25**

Ficha técnica:

Autoras: Cláudia Silva, Rita Guerra, Paula Carvalho.
Design gráfico e ilustrações: ROD Saturnino

Como citar:

Silva, C., Guerra, R., & Carvalho, P. (2024). Desmascarar o Discurso de Ódio: O Conhecimento é a nossa arma. ISBN: 978-989-781-821-9.

SOBRE O QUE É ESTA BROCHURA?

Esta brochura tem como objetivo informar o público sobre o discurso de ódio online na sociedade portuguesa e apresentar os principais resultados do projeto kNOwHATE (ver <https://knowhate.eu/pt-pt/>).

O projeto kNOwHATE, financiado pela UE, desenvolveu uma abordagem integrada e interdisciplinar para compreender o discurso de ódio online. Esta abordagem foi utilizada para analisar dados textuais de redes sociais, incluindo tweets e comentários a vídeos do YouTube, escritos em português europeu. Estes dados foram explorados para o desenvolvimento de modelos de deteção automática de discurso de ódio online.

Para destacar as características psicossociais e linguísticas do discurso de ódio online em português, apresentamos nesta brochura vários exemplos de dados que foram analisados neste projeto. Esperamos que este recurso seja útil para educadores, especialistas em literacia mediática, legisladores e para a sociedade civil, em geral, ajudando-os a compreender a natureza complexa e multifacetada do discurso de ódio, incluindo as formas indiretas e subtis de discurso de ódio, que são muitas vezes difíceis de detetar e combater.

Embora esta brochura seja um recurso independente, também faz parte de uma campanha mais alargada contra o discurso de ódio online em Portugal. Esta campanha inclui uma série de podcasts e um breve vídeo, especificamente criado para divulgação nas redes sociais.

Para mais informações, procurem a hashtag **#CortarOMalPelaRaiz** no Instagram.

Juntem-se a nós nesta campanha e espalhem a palavra contra o ódio!

AVISO

Esta brochura inclui exemplos que contêm expressões diretas e indiretas de ódio e linguagem depreciativa e ofensiva. Procuramos fazer um equilíbrio entre a reprodução de conteúdos de ódio e a necessidade de sensibilização e o conhecimento sobre esta perigosa forma de discurso. Utilizamos mensagens reais de discurso de ódio retiradas de dados publicamente disponíveis nas redes sociais, de modo a minimizar o risco de danos e a seguir o Código de Conduta da APA para a investigação, que recomenda que "o risco de danos não deve ser maior do que na vida normal, ou seja, os indivíduos não devem ser expostos a riscos maiores ou adicionais aos encontrados nos seus estilos de vida normais". Expor as pessoas a discursos de ódio pode desencadear consequências negativas. Deixamos várias fontes onde podem encontrar informações sobre o impacto negativo do discurso de ódio para vítimas e espectadores, bem como um conjunto de estratégias para o combater:

1. <https://www.coe.int>
2. <https://unesdoc.unesco.org>
3. <https://www.undocs.org>



O QUE É O DISCURSO DE ÓDIO ONLINE?

Não existe uma definição académica clara e consensual de discurso de ódio online, o que coloca desafios para a sua compreensão e combate. A definição adotada no kNOwHATE baseia-se na literatura da Psicologia Social sobre preconceito e discriminação (Allport, 1954), em revisões integrativas recentes (Siegel, 2020), bem como nas orientações fornecidas pelo Conselho da Europa na sua última Recomendação sobre discurso de ódio (CM/Rec/2022/16).

Especificamente, o kNOwHATE aborda o discurso de ódio online como um fenómeno intergrupar. O discurso de ódio tem como alvos grupos sociais ou indivíduos que são percebidos como pertencendo a esses grupos, sendo atacados pelas características associadas a esses grupos e não necessariamente devido a um comportamento ou ação concreta.

o discurso de ódio online manifesta-se através do uso de linguagem depreciativa, motivada por preconceitos, que espalha, incita, promove ou justifica o ódio, a exclusão e/ou a violência/agressão, contra grupos ou indivíduos com base na perceção de pertença a um grupo (por exemplo, características percebidas como etnia, raça, orientação sexual, etc.).

O discurso de ódio online ocorre na esfera digital, principalmente em aplicações e plataformas de redes sociais, mas também em jogos online, aplicações de realidade virtual e outras plataformas dos media digitais.



SABIAM QUE EXISTEM DIFERENTES TIPOS DE DISCURSO DE ÓDIO ONLINE?

No kNOwHATE, consideramos duas formas de discurso de ódio: direta e indireta. Ambas as formas envolvem a expressão de ódio em que o orador espalha, incita, promove ou justifica o ódio, a exclusão e/ou a violência contra um grupo ou pessoa-alvo com base na sua pertença a um grupo. No entanto, fazem-no de forma diferente:



DISCURSO DE ÓDIO DIRETO

manifesta-se através de um discurso explícito, declaradamente tendencioso e inflamatório, que recorre frequentemente a insultos e linguagem depreciativa.

Racismo o c@ralho! se não fossem esses parasitas da sociedade que não querem fazer nada, Portugal era um paraíso.



DISCURSO DE ODIO INDIRETO

manifesta-se de forma indireta ou subtil, e o seu significado está implícito no texto, tendo de ser inferido pelo leitor. Nestas mensagens, é comum o uso de várias estratégias retóricas e discursivas, incluindo, por exemplo, a ironia, o sarcasmo, o humor, os eufemismos e as perguntas retóricas.

Coitadinhos dos “feirantes”, vão ficar sem os benefícios.



QUE TIPO DE DISCURSO DE ÓDIO ONLINE É MAIS COMUM EM PORTUGAL?

Os nossos resultados mostram que o discurso de ódio indireto prevalece sobre o discurso de ódio direto, independentemente do grupo-alvo ou das redes sociais utilizadas. Estes resultados estão em linha com investigações anteriores realizadas em Portugal (e.g., Carvalho et al. 2023; Silva & Carvalho, 2023) e noutros contextos culturais (e.g., Baider, 2023).



COMO IDENTIFICAR O DISCURSO DE ÓDIO ONLINE?

Embora seja relativamente fácil identificar formas de discurso de ódio direto, nomeadamente devido à utilização de linguagem depreciativa explícita (por exemplo, insultos), o mesmo não se aplica ao discurso de ódio indireto. Neste caso, é preciso ler nas entrelinhas e identificar estratégias linguístico-discursivas que mobilizam o ódio. A identificação dessas estratégias permite não só desvendar o discurso de ódio, mas também produzir um contra-discurso eficaz para o seu combate.

Há estratégias comuns mobilizadas tanto no discurso de ódio direto como no indireto, embora a sua prevalência possa variar consoante o tipo de discurso. Entre outras, destacam-se os estereótipos negativos, a desumanização, a ameaça ou expressão de emoções negativas, como o ódio e a raiva. Existem também estratégias discursivas distintivas no discurso de ódio indireto, tais como a negação do ódio ou a inversão de papéis. Eis alguns exemplos :



ESTEREÓTIPOS NEGATIVOS

1

Os estereótipos negativos (ou seja, características negativas e inexatas associadas aos grupos sociais visados ou aos seus membros) são frequentemente empregues para denegrir ou humilhar o alvo através de generalizações negativas falaciosas.

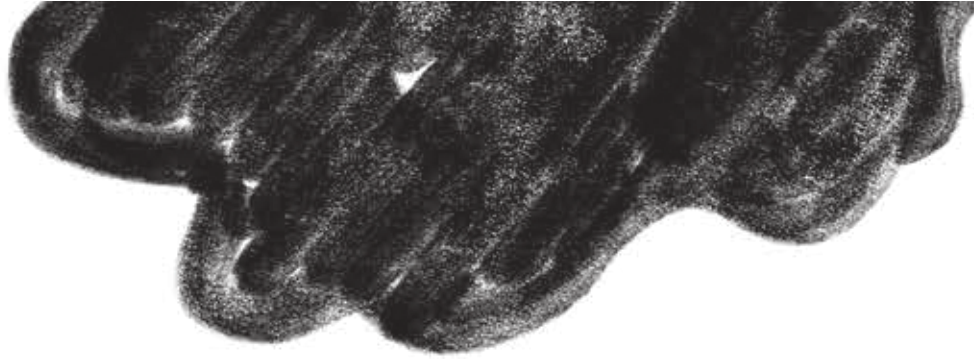
Coitad@. Vcs deitam se com primos, tios, irmaos etc praticam incest# entre familia que Casam crianças com adultos e forçam nas a engravidar

DESUMANIZAÇÃO

2

A desumanização envolve encarar os grupos-alvo (por exemplo, pessoas racializadas) como não-humanos, facilitando ações que causem dano nos mesmos. É frequentemente materializada através de comparações e metáforas que reduzem os grupos-alvo a animais, objetos ou autómatos.

Ciganos são como javalis, são animais selvagens



AMEAÇAS REALISTAS E SIMBÓLICAS

O discurso de ódio envolve frequentemente perceções e afirmações de que o grupo-alvo representa uma ameaça para o endogrupo (i.e., grupo de pertença), seja ela realista (ao poder, recursos, bem-estar geral, saúde ou segurança do grupo) ou simbólica (à religião, valores, sistema de crenças, ideologia, moralidade ou visão do mundo).



REALISTA

Quando os ciganos assaltam, matam ou agridem as forças de segurança não se diz ciganos, diz-se jovens, quando é para elogiar já se diz ciganos

SIMBÓLICA

Sinto-me ofendido quando a bandeira do meu país é "transvirada" alterada e usada para uma causa "esquerdista" Vão para a cama com quem quiserem, mas o respeito e a ordem nos nossos símbolos não podem ser adulterados ou alienados, querem respeito...tenham respeito!!!

ÓDIO

O ódio é uma emoção negativa poderosa, motivada por avaliações de que os outros têm intenções malévolas de causar danos ao endogrupo (i.e., grupo de pertença). O ódio prevê ações específicas como causar danos no grupo-alvo e potencialmente até o aniquilar, quer a nível psicológico (e.g., humilhação), social (e.g., exclusão, negligência) ou físico (e.g., morte, tortura, agressão).



4

“MATAR O CARALHO..VAI PRÁ TUA
TERRA.MAS AFINAL DE QUE TRIBO
ÉS..MAMADOU .E NÃO HÁ QUÊ
LHE.TRATE DA SAÚDINHA.”

RAIVA

A raiva é sentida quando as ações do grupo-alvo são vistas como injustas ou desviantes das normas aceitáveis do grupo de pertença. Exprime uma motivação para remover obstáculos, para mudar os outros, através do confronto, crítica ou ataque.



5

“Força Mário, fizeste muito bem. Essa escumalha vai ter que aprender a respeitar os Portugueses e parar de gozar com os cidadãos de bem. Viva Portugal”

“Agora quero ver quem vem defender os pobres coitadinhos desta ciganada. Tudo podre neste nosso país onde estes marginais são defendidos em praca pública por esses porcos bloquistas comunistas.”

NEGAÇÃO DO ÓDIO

6

A negação do ódio é uma estratégia que tem como objetivo apresentar positivamente o endogrupo (i.e. grupo de pertença), em que o orador procura proteger os seus argumentos de acusações de ódio, negando-o para manter a credibilidade, reduzir a perceção de ódio e preservar a legitimidade do seu argumento.

"Omg eu juro, nao sou racista mas se tivesse de tirar uma raça da Europa seria esta"

INVERSÃO DE PAPÉIS

7

A inversão de papéis tem como objetivo obscurecer as desigualdades sistémicas, invertendo dinâmicas de poder e minando a opressão. Ocorre quando um membro de um grupo social privilegiado ou dominante (por exemplo, uma pessoa de classe alta, branca heteronormativa) afirma ser vítima de discriminação ou preconceito enquanto retrata outros grupos (por exemplo, grupos racializados) como opressores que representam uma ameaça para o seu grupo de pertença (i.e. endogrupo).

“SE OS PORTUGUESES NAO VOTAR EM ANDRÉ VENTURA ,VAI CHEGAR O DIA QUE NAO PODEMOS SAIR À RUA!... VEJAM O QUE SE PASSA COM AUTORIDADES .,E AINDA ELES ESTÃO ARMADOS E COM A FARDA DE SER AUTORIDADE !!!”

Dica: Estas estratégias são muitas vezes complexas, uma vez que empregam dispositivos retóricos como a ironia, sarcasmo, humor e figuras de linguagem, como as metáforas e hipérbolos. Fiquem atentos/as.





O QUE É O CONTRADISCURSO ONLINE?

O contradiscurso online é uma resposta direta ao discurso de ódio com o objetivo de o combater. Esta resposta pode ser organizada (por exemplo, hashtags #ichbinhier, #brasileirasnaosecalam) ou espontânea. O contra-discurso online pode basear-se em texto ou imagens e a investigação tem identificado algumas características do seu sucesso para alterar o discurso original para um tom mais favorável: por exemplo, apresentação de factos para corrigir afirmações ou perceções erradas; aviso de possíveis consequências online e offline; uso de emoções positivas, como a empatia; uso do humor. No entanto, é importante notar que a eficácia destas estratégias pode variar consoante o grupo-alvo.

No kNOWHATE, com base nesta investigação anterior e em estratégias psicossociais bem estabelecidas para reduzir a discriminação (por exemplo, identidades inclusivas), analisámos estratégias de contra-discurso que envolvem técnicas discursivas baseadas em contraestereótipos, empatia, identidades inclusivas e apelos a normas legais e sociais.

CONTRAESTEREÓTIPOS

Referem-se a representações que desafiam ou contrariam crenças e perceções culturais amplamente atribuídas a um grupo social. Salientam informação incongruente com estereótipos prevalentes, desafiando preconceitos e promovendo perceções mais exatas dos grupos sociais.



“Olá já procuraste alguma vez te informar com factos verídicos acerca dos subsídios declarados para a comunidade cigana? Penso que não mas tens vários sites que o podem provar a tua gigantesca burrice disfarçada de preconceito étnico. Se precisares até eu posso te mandar alguns”



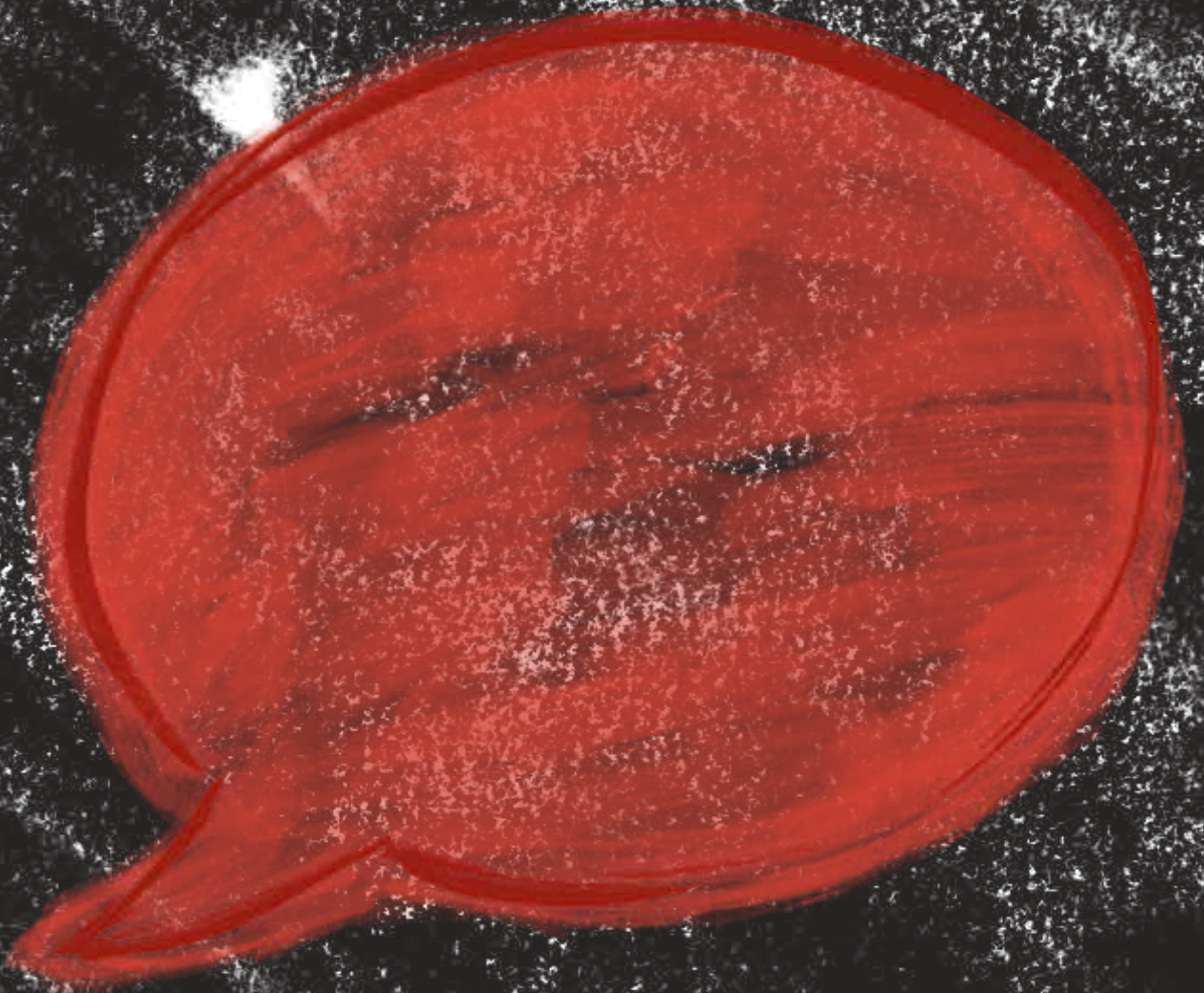
EMPATIA

envolve tanto aspetos afetivos (correspondência de emoções e preocupação empática) como cognitivos (tomada de perspetiva), e está associada a um aumento da tolerância e de comportamentos de ajuda.

“**devias apanhar um bofia que te pedisse para baixar as calças.** abuso de autoridade tb é crime. sobre qual acusação estão o estão a deter? nao uso de máscara?!”

“gostava de saber se não fossem os estrangeiros a trabalhar nas obras, a cuidar dos nossos idosos, a trabalhar no comercio e hotelaria se os portugueses que não aceitam trabalhar por esses salários o que íamos fazer!!! falta de noção da realidade”





IDENTIDADES INCLUSIVAS

Esta estratégia envolve a ativação de identidades comuns, levando as pessoas a mudar a sua categorização de entidades separadas ("nós" vs "eles") para um grupo mais inclusivo e abrangente ("nós").

"Somos a raça humana e temos de saber viver juntos neste mundo. Em vez de nos focarmos nas nossas diferenças, **vamo-nos focar no que temos em Comum**".

"Certíssimo. Os portugueses têm de se unir. Sejamos brancos, pretos, ciganos ou o quer que seja, **temos uma coisa em comum: somos portugueses**".

NORMAS LEGAIS E SOCIAIS

O apelo às normas legais funciona como um dissuasor do discurso de ódio online, ao incutir o medo de punição. Estas regulamentações também estabelecem uma norma social que considera a linguagem de ódio socialmente indesejável e inadequada, restringindo o envolvimento no discurso de ódio e promovendo esforços coletivos contra o mesmo.

"Quem nasce em Portugal é português. Não interessa se é etnicamente europeu, africano, americano, asiático etc..."

"O mundo não é só 0 e 1, neste caso XX e XY, também existe os XXY. Também têm direito à vida, ou será que não?"

"devias apanhar um bofia que te pedisse para baixar as calças. abuso de autoridade tb é crime. sobre qual acusação estão o estão a deter? nao uso de máscara?!"



COMO PODEM COMBATER O DISCURSO DE ÓDIO?

Existem várias formas de combater o discurso de ódio online, que vão desde a moderação de conteúdos, à produção de contra-discursos e contra-narrativas até à aplicação de medidas legais. Com base na investigação anterior e nas conclusões do kNOwHATE, destacamos potenciais mecanismos para combater o discurso de ódio online.

Denunciá-lo nos canais das redes sociais, marcando os comentários como potencialmente odiosos. As plataformas de redes sociais podem ter diferentes formas de denunciar conteúdos de ódio. Estes comentários podem ser denunciados nos formulários disponibilizados nas plataformas para este fim (e.g., o Facebook permite denunciar comentários). Mesmo que o resultado da denúncia possa não ser o esperado, as queixas formais através dos canais oficiais são importantes no combate ao discurso de ódio.

Utilizar os media digitais para criar contra-narrativas positivas, capacitando assim as comunidades-alvo. Em Portugal, destacam-se várias iniciativas, como o podcast "Mitos e Verdades", que contraria estereótipos sobre a comunidade cigana, ou a campanha #brasileirasnaosecalam, que partilha histórias e apoia mulheres brasileiras que enfrentam xenofobia. O canal do YouTube "Dicionário Interseccional" educa sobre conceitos de discurso de ódio. A campanha #ABCLGBTQIA+ da TV Foxlife Portugal e da ILGA aborda questões LGBTQIA+.

Os resultados da nossa análise, baseada em 24.739 comentários a 88 vídeos do YouTube e 29.846 tweets extraídos de 2775 conversas no Twitter/X, mostraram que as estratégias mais prevalentes nas mensagens de contra-discurso foram a utilização de contra-estereótipos, a empatia para com as comunidades visadas, bem como a mobilização de normas sociais e identidades inclusivas. Embora o kNOwHATE não tenha testado a eficácia destas estratégias para travar especificamente o discurso de ódio, existem evidências sólidas de investigação anterior que mostram que, entre outras, estas estratégias são eficazes para reduzir a discriminação e o preconceito.

COMUNIDADES RACIALIZADAS SOS RACISMO

O QUE CARACTERIZA O DDO CONTRA ESTAS COMUNIDADES?

Segundo a SOS Racismo, as comunidades cigana e afro-descendente portuguesas são as minorias étnicas mais visadas pelo discurso de ódio em Portugal. O discurso de ódio racista baseia-se na crença de uma hierarquia de raças que oprime grupos minoritários, desumanizando-os, e justificando estereótipos negativos e comportamentos discriminatórios. Isto manifesta-se, subtil e explicitamente, em em várias esferas públicas e privadas, incluindo relações interpessoais, políticas, sociais e discursos institucionais do Estado, forças de segurança, tribunais, escolas, serviços de saúde, transportes habitação, desporto e imprensa.



COMUNIDADES LGBTI+ ILGA

O QUE CARACTERIZA O DÓIO CONTRA ESTAS COMUNIDADES?

O discurso de ódio dirigido à comunidade LGBTI+ baseia-se frequentemente em argumentos biológicos/essencialistas ou religiosos, retratando as pessoas LGBTI+ como "anormais" ou "amorais". Isto pode visar aspetos específicos da orientação sexual, identidade de género ou características sexuais. Além disso, o discurso de ódio associa frequentemente as pessoas LGBTI+ a perturbações psicológicas, perpetuando a falsa noção de que têm problemas de saúde mental. Uma característica comum deste discurso de ódio é a ridicularização das identidades, como por exemplo responder a uma pessoa não-binária com "Oh, Eu identifico-me como um helicóptero". Outra característica frequente é rotular o seu comportamento como um desvio das normas sociais ou "antinatural".

ILGA

INTERVENÇÃO LÉSBICA, GAY,
BISSEXUAL, TRANS E INTERSEXO

COMUNIDADES MIGRANTES CASA DO BRASIL EM LISBOA

O QUE CARACTERIZA O DOO CONTRA ESTAS COMUNIDADES?

De acordo com o projeto "#MigraMyths - Desmistificando a Imigração", liderado pela Casa do Brasil de Lisboa, inquéritos realizados em 2020 e 2021 revelaram que 75% a 85% dos 240 participantes imigrantes sofreram discriminação baseada em mitos e estereótipos negativos em Portugal. As mulheres foram a maioria em ambos os inquéritos, em linha com os estereótipos sobre as mulheres brasileiras em Portugal, relacionados à prostituição e "roubar maridos". Estes estereótipos negativos têm um impacto significativo nas suas vidas, levando à discriminação diária na habitação, assédio sexual e desqualificação profissional. Os participantes mencionaram ainda outros preconceitos, como a associação de imigrantes à criminalidade e ao "roubo de empregos".



Hugging Face

TESTEM O PROTÓTIPO

Com o objetivo de desenvolver modelos computacionais para um sistema de deteção automática de discurso de ódio online, o kNOwHATE recolheu dados do YouTube e do Twitter, incluindo 24.739 comentários de 88 vídeos do YouTube e 29.846 tweets de 2.775 conversas no Twitter/X. Estes dados foram meticulosamente anotados por uma equipa de anotadores treinados, incluindo psicólogos sociais e linguistas.

Convidamos os leitores a explorar o nosso conjunto de dados e a simular como um sistema de deteção automática de discurso de ódio online poderia funcionar aqui: <https://huggingface.co/knowhate>

Este protótipo não é a ferramenta final e pode conter erros e imprecisões na identificação do discurso de ódio.

Para o experimentarem, escrevam uma frase ou expressão (por exemplo, "volta para a tua terra"/ volta para o teu país"), insiram-na na caixa e cliquem em "submeter". O protótipo informará se o texto introduzido contém ou não discurso de ódio.

REFERÊNCIAS

Esta brochura foi elaborada principalmente a partir do relatório entregue pelo grupo de trabalho 2 do kNOwHATE, bem como das diretivas de anotação fornecidas pelo grupo de trabalho 3. Abaixo encontram-se algumas referências importantes utilizadas neste documento.

Allport, G. W. (1954). *The Nature of Prejudice*. Addison-Wesley

Baider, F., & Constantinou, M. (2020). Covert hate speech: A contrastive study of Greek and Greek Cypriot online discussions with an emphasis on irony. *Journal of Language Aggression and Conflict*, 8(2), 262-287.

Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social issues and policy review*, 3(1), 141-177.

Benesch, Susan; Derek Ruths, Kelly P. Dillon, Haji Mohammad Saleem, and Lucas Wright. (2016). "Counterspeech on Twitter: A Field Study." *Dangerous Speech Project*. Available at: <https://dangerousspeech.org/counterspeech-on-twitter-a-field-study>

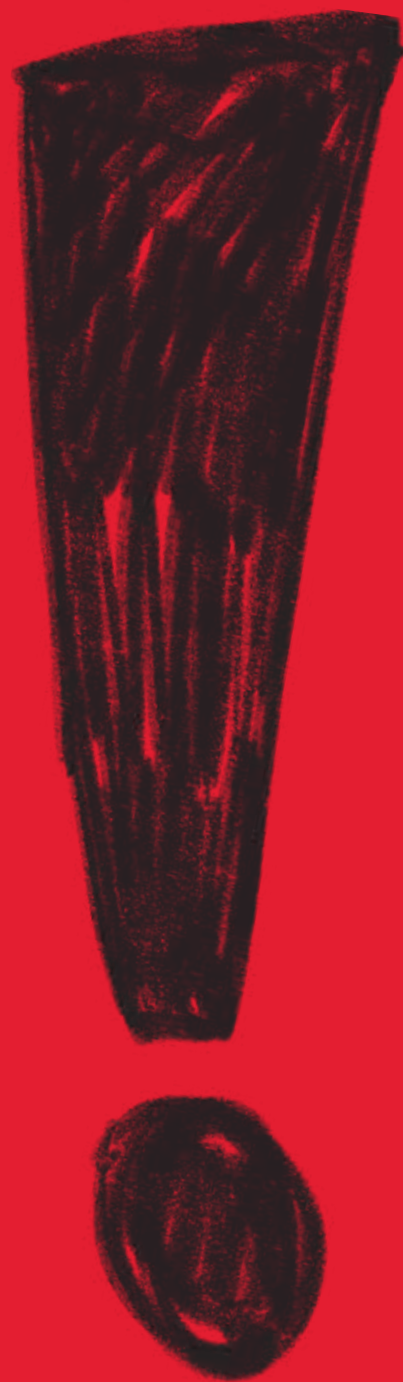
Carvalho, P., Caled, D., Silva, C., Batista, F., & Ribeiro, R. (2023). The expression of hate speech against Afro-descendant, Roma, and LGBTQ+ communities in YouTube comments. *Journal of Language Aggression and Conflict*. <https://doi.org/10.1075/jlac.00085.car>

Fischer, A., Halperin, E., Canetti, D., & Jasini, A. (2018). Why we hate. *Emotion Review*, 10(4), 309-320.

Siegel, A. A. (2020). Online hate speech. In N. Persily and J. A. Tucker (Eds), *Social media and democracy: The state of the field, prospects for reform* (pp. 56-88). Cambridge University Press.

Silva, C., & Carvalho, P. (2023). When can compliments and humour be considered hate speech? A perspective from target groups in Portugal. *Comunicação e Sociedade*.

DOI: [https://doi.org/10.17231/comsoc.43\(2023\).4135](https://doi.org/10.17231/comsoc.43(2023).4135)



Unidades de Investigação

bru _iscte

Business Research Unit

cis _iscte

Centro de Investigação
e Intervenção Social

istar _iscte

Centro de Investigação
em Ciências da Informação,
Tecnologias e Arquitetura

cies _iscte

Centro de Investigação
e Estudos de Sociologia



**SOCIODIGITAL LAB
FOR PUBLIC POLICY**
Associate Laboratory

INTERACTIVE
TECHNOLOGIES
INSTITUTE



Membros do Consórcio KNOwHATE



CICDR
COMISSÃO PARA A IGUALDADE
E CONTRA A DISCRIMINAÇÃO RACIAL



**25
ANOS**



<https://knowhate.eu>

**KNOWING
ONLINE
HATE SPEECH**

KNOWLEDGE + AWARENESS = TACKLING HATE



Financiado pela União Europeia. As visões e opiniões expressas são, no entanto, apenas do(s) autor(es) e não refletem necessariamente as da União Europeia ou do Projeto kNOwHATE. Nem a União Europeia nem o Projeto kNOwHATE podem ser responsabilizados por eles.