

## An overview of hate speech datasets

Project	Language	Source	Size (no. of posts)	Tags	Reference
Are They our Brothers? Analysis and Detection of Religious Hate Speech in the Arabic Twittersphere	Arabic	Twitter	6,136	Binary – Hate/Not hate	Albadi, N., Kurdi, M. and Mishra, S., 2018. Are they Our Brothers? Analysis and Detection of Religious Hate Speech in the Arabic Twittersphere. In: International Conference on Advances in Social Networks Analysis and Mining. Barcelona, Spain: IEEE, pp.69-76.
Multilingual and Multi-Aspect Hate Speech Analysis	Arabic	Twitter	3,353	Hostility, Directness, Target Attribute, Target Group, How annotators felt on seeing the tweet	Ousidhoum, N., Lin, Z., Zhang, H., Song, Y. and Yeung, D., 2019. Multilingual and Multi-Aspect Hate Speech Analysis. ArXiv
L-HSAB: A Levantine Twitter Dataset for Hate Speech and Abusive Language	Arabic	Twitter	5,846	Ternary (Hate, Abusive, Normal)	Mulki, H., Haddad, H., Bechikh, C. and Alshabani, H., 2019. L-HSAB: A Levantine Twitter Dataset for Hate Speech and Abusive Language. In: Proceedings of the Third Workshop on Abusive Language Online. Florence, Italy: Association for Computational Linguistics, pp.111-118.
Abusive Language Detection on Arabic Social Media	Arabic	Twitter	1,100	Ternary (Obscene, Offensive but not obscene, Clean)	Mubarak, H., Darwish, K. and Magdy, W., 2017. Abusive Language Detection on Arabic Social Media. In: Proceedings of the First Workshop on Abusive Language Online. Vancouver, Canada: Association for Computational Linguistics, pp.52-56.
Abusive Language Detection on Arabic Social Media	Arabic	Posts (Al Jazeera platform)	32,000	Ternary (Obscene, Offensive but not obscene, Clean)	Mubarak, H., Darwish, K. and Magdy, W., 2017. Abusive Language Detection on Arabic Social Media. In: Proceedings of the First Workshop on Abusive Language Online. Vancouver, Canada: Association for Computational Linguistics, pp.52-56.
Dataset Construction for the Detection of Anti-Social	Arabic	Youtube posts	15,050	Binary (Offensive, Not)	Alakrot, A., Murray, L. and Nikolov, N., 2018. Dataset Construction for the Detection of Anti-Social

Behaviour in Online Communication in Arabic					Behaviour in Online Communication in Arabic. <i>Procedia Computer Science</i> , 142, pp.174-181.
Datasets of Slovene and Croatian Moderated News Comments	Croatian	24sata website, posts	17,000,000	Binary (Deleted, Not)	Ljubešić, N., Erjavec, T. and Fišer, D., 2018. Datasets of Slovene and Croatian Moderated News Comments. In: <i>Proceedings of the 2nd Workshop on Abusive Language Online (ALW2)</i> . Brussels, Belgium: Association for Computational Linguistics, pp.124-131.
Offensive Language and Hate Speech Detection for Danish	Danish	Twitter, Reddit, newspaper comments	3,600	Binary (Offensive, Not), Within Offensive (Target, Not), Within Target (Individual, Group, Other)	Sigurbergsson, G. and Derczynski, L., 2019. Offensive Language and Hate Speech Detection for Danish. <i>ArXiv</i> .
Automated Hate Speech Detection and the Problem of Offensive Language	English	Twitter	24,802	Hierarchy (Hate, Offensive, Neither)	Davidson, T., Warmsley, D., Macy, M. and Weber, I., 2017. Automated Hate Speech Detection and the Problem of Offensive Language. <i>ArXiv</i> .
Hate Speech Dataset from a White Supremacy Forum	English	Stormfront (Forum)	9,916	Ternary (Hate, Relation, Not)	de Gibert, O., Perez, N., García-Pablos, A., and Cuadros, M., 2018. Hate Speech Dataset from a White Supremacy Forum. In: <i>Proceedings of the 2nd Workshop on Abusive Language Online (ALW2)</i> . Brussels, Belgium: Association for Computational Linguistics, pp.11-20.
Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter	English	Twitter	16,914	3-topic (Sexist, Racist, Not)	Waseem, Z. and Horvy, D., 2016. Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. In: <i>Proceedings of the NAACL Student Research Workshop</i> . San Diego, California: Association for Computational Linguistics, pp.88-93.
Detecting Online Hate Speech Using Context Aware Models	English	FoxNews, posts	1,528	Binary (Hate / not)	Gao, L. and Huang, R., 2018. Detecting Online Hate Speech Using Context Aware Models. <i>ArXiv</i> .
Are You a Racist or Am I Seeing Things? Annotator Influence on Hate Speech	English	Twitter	4,033	Multi-topic (Sexist, Racist, Neither, Both)	Waseem, Z., 2016. Are You a Racist or Am I Seeing Things? Annotator Influence on Hate Speech Detection on Twitter. In: <i>Proceedings of 2016</i>

Detection on Twitter					EMNLP Workshop on Natural Language Processing and Computational Social Science. Copenhagen, Denmark: Association for Computational Linguistics, pp.138-142.
When Does a Compliment Become Sexist? Analysis and Classification of Ambivalent Sexism Using Twitter Data	English	Twitter	712	Hierarchy of Sexism (Benevolent sexism, Hostile sexism, None)	Jha, A. and Mamidi, R., 2017. When does a Compliment become Sexist? Analysis and Classification of Ambivalent Sexism using Twitter Data. In: Proceedings of the Second Workshop on Natural Language Processing and Computational Social Science. Vancouver, Canada: Association for Computational Linguistics, pp.7-16.
Overview of the Task on Automatic Misogyny Identification at IberEval 2018	English	Twitter	3,977	Binary (misogyny / not), 5 categories (stereotype, dominance, derailing, sexual harassment, discredit), target of misogyny (active or passive)	Fersini, E., Rosso, P. and Anzovino, M., 2018. Overview of the Task on Automatic Misogyny Identification at IberEval 2018. In: Proceedings of the Third Workshop on Evaluation of Human Language Technologies for Iberian Languages (IberEval 2018).
CONAN - COunter NArratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech	English	Synthetic / Facebook posts	1,288	Binary (Islamophobic / not), multi-topic (Culture, Economics, Crimes, Rapism, Terrorism, Women Oppression, History, Other/generic)	Chung, Y., Kuzmenko, E., Tekiroglu, S. and Guerini, M., 2019. CONAN - COunter NArratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, pp.2819-2829.
Characterizing and Detecting Hateful Users on Twitter	English	Twitter	4,972	Binary (hateful/not)	Ribeiro, M., Calais, P., Santos, Y., Almeida, V. and Meira, W., 2018. Characterizing and Detecting Hateful Users on Twitter. ArXiv
A Benchmark Dataset for Learning to Intervene in Online Hate	English	Platform Gab, posts	33,776	Binary (hateful/not)	Qian, J., Bethke, A., Belding, E. and Yang Wang, W., 2019. A Benchmark Dataset for Learning to Intervene in

Speech					Online Hate Speech. ArXiv
A Benchmark Dataset for Learning to Intervene in Online Hate Speech	English	Reddit	22,324	Binary (hateful/not)	Qian, J., Bethke, A., Belding, E. and Yang Wang, W., 2019. A Benchmark Dataset for Learning to Intervene in Online Hate Speech. ArXiv
Multilingual and Multi-Aspect Hate Speech Analysis	English	Twitter	5,647	Hostility, Directness, Target attribute and Target group.	Ousidhoum, N., Lin, Z., Zhang, H., Song, Y. and Yeung, D., 2019. Multilingual and Multi-Aspect Hate Speech Analysis. ArXiv
Exploring Hate Speech Detection in Multimodal Publications	English	Twitter	149,823	Six primary categories (No attacks to any community, Racist, Sexist, Homophobic, Religion based attack, Attack to other community)	Gomez, R., Gibert, J., Gomez, L. and Karatzas, D., 2019. Exploring Hate Speech Detection in Multimodal Publications. ArXiv
Predicting the Type and Target of Offensive Posts in Social Media	English	Twitter	14,100	Branching structure of tasks: Binary (Offensive, Not), Within Offensive (Target, Not), Within Target (Individual, Group, Other)	Zampieri, M., Malmasi, S., Nakov, P., Rosenthal, S., Farra, N. and Kumar, R., 2019. SemEval-2019 Task 6: Identifying and Categorizing Offensive Language in Social Media (OffensEval). ArXiv,.
hatEval, SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter	English	Twitter	13,000	Branching structure of tasks: Binary (Hate, Not), Within Hate (Group, Individual), Within Hate (Agressive, Not)	Basile, V., Bosco, C., Fersini, E., Nozza, D., Patti, V., Pardo, F., Rosso, P. and Sanguinetti, M., 2019. SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter. In: Proceedings of the 13th International Workshop on Semantic Evaluation. Minneapolis, Minnesota: Association for Computational Linguistics, pp.54-63.
Peer to Peer Hate: Hate Speech Instigators and Their Targets	English	Twitter	27,330	Binary (Hate/Not), only for tweets which have both a Hate Instigator and Hate Target	ElSherief, M., Nilizadeh, S., Nguyen, D., Vigna, G. and Belding, E., 2018. Peer to Peer Hate: Hate Speech Instigators and Their Targets. In: Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018). Santa Barbara, California: University of California, pp.52-61.
Overview of the HASOC track at FIRE 2019: Hate Speech and Offensive Content	English	Twitter and Facebook	7,005	Branching structure of tasks. A: Hate / Offensive or Neither, B: Hatespeech, Offensive, or Profane, C: Targeted or Untargeted	Mandl, T., Modha, S., Majumder, P., Patel, D., Dave, M., Mandlia, C. and Patel, A., 2019. Overview of the HASOC track at FIRE 2019. In: Proceedings

Identification in Indo-European Languages					of the 11th Forum for Information Retrieval Evaluation
Large Scale Crowdsourcing and Characterization of Twitter Abusive Behavior	English	Twitter	80,000	Multi-thematic (Abusive, Hateful, Normal, Spam)	Founta, A., Djouvas, C., Chatzakou, D., Leontiadis, I., Blackburn, J., Stringhini, G., Vakali, A., Sirivianos, M. and Kourtellis, N., 2018. Large Scale Crowdsourcing and Characterization of Twitter Abusive Behavior. ArXiv
A Large Labeled Corpus for Online Harassment Research	English	Twitter	35,000	Binary (Harassment, Not)	Golbeck, J., Ashktorab, Z., Banjo, R., Berlinger, A., Bhagwan, S., Buntain, C., Cheakalos, P., Geller, A., Gergory, Q., Gnanasekaran, R., Gnanasekaran, R., Hoffman, K., Hottle, J., Jienjiltert, V., Khare, S., Lau, R., Martindale, M., Naik, S., Nixon, H., Ramachandran, P., Rogers, K., Rogers, L., Sarin, M., Shahane, G., Thanki, J., Vengataraman, P., Wan, Z. and Wu, D., 2017. A Large Labeled Corpus for Online Harassment Research. In: Proceedings of the 2017 ACM on Web Science Conference. New York: Association for Computing Machinery, pp.229-233.
Ex Machina: Personal Attacks Seen at Scale, Personal attacks	English	Wikipedia posts	115,737	Binary (Personal attack, Not)	Wulczyn, E., Thain, N. and Dixon, L., 2017. Ex Machina: Personal Attacks Seen at Scale. ArXiv
Ex Machina: Personal Attacks Seen at Scale, Toxicity	English	Wikipedia posts	100,000	Toxicity/healthiness judgement (very toxic, neutral, very healthy)	Wulczyn, E., Thain, N. and Dixon, L., 2017. Ex Machina: Personal Attacks Seen at Scale. ArXiv
Detecting cyberbullying in online communities	English	World of Warcraft, posts	16,975	Binary (Harassment, Not)	Bretschneider, U. and Peters, R., 2016. Detecting Cyberbullying in Online Communities. Research Papers, 61.
Detecting cyberbullying in online communities	English	League of Legends, posts	17,354	Binary (Harassment, Not)	Bretschneider, U. and Peters, R., 2016. Detecting Cyberbullying in Online Communities. Research Papers, 61.
A Quality Type-aware Annotated Corpus and Lexicon for Harassment Research	English	Twitter	24,189	Multi-topic harassment detection	Rezvan, M., Shekarpour, S., Balasuriya, L., Thirunarayan, K., Shalin, V. and Sheth, A., 2018. A Quality Type-aware Annotated Corpus and Lexicon for Harassment Research. ArXiv
Ex Machina:	English	Wikipedia	160,000	Aggression/friendliness	Wulczyn, E., Thain, N. and Dixon,

Personal Attacks Seen at Scale, Aggression and Friendliness		posts		judgement on a 5 point scale. (very aggressive, neutral, very friendly).	L., 2017. Ex Machina: Personal Attacks Seen at Scale. ArXiv
CONAN - COunter NArratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech	French	Synthetic / Facebook	1,719	Binary (Islamophobic / not), Multi-topic (Culture, Economics, Crimes, Rapism, Terrorism, Women Oppression, History, Other/generic)	Chung, Y., Kuzmenko, E., Tekiroglu, S. and Guerini, M., 2019. CONAN - COunter NArratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, pp.2819-2829.
Multilingual and Multi-Aspect Hate Speech Analysis	French	Twitter	4,014	Detailed taxonomy with cross-cutting attributes: Hostility, Directness, Target Attribute, Target Group, How annotators felt on seeing the tweet. Details of task: Gender, Sexual orientation, Religion, Disability	Ousidhoum, N., Lin, Z., Zhang, H., Song, Y. and Yeung, D., 2019. Multilingual and Multi-Aspect Hate Speech Analysis. ArXiv
Measuring the Reliability of Hate Speech Annotations: The Case of the European Refugee Crisis	German	Twitter	469	Binary (Anti-refugee hate, None)	Ross, B., Rist, M., Carbonell, G., Cabrera, B., Kurowsky, N. and Wojatzki, M., 2017. Measuring the Reliability of Hate Speech Annotations: The Case of the European Refugee Crisis. ArXiv
Detecting Offensive Statements Towards Foreigners in Social Media	German	Facebook	5,836	Hierarchical (Anti-foreigner prejudice, split into (1) slightly offensive/offensive and (2) explicitly/substantially offensive). 6 targets (Foreigner, Government, Press, Community, Other, Unknown)	Bretschneider, U. and Peters, R., 2017. Detecting Offensive Statements towards Foreigners in Social Media. In: Proceedings of the 50th Hawaii International Conference on System Sciences.
GermEval 2018	German	Twitter	8,541	Branching structure: Binary (Offense, Other), 3 levels within Offense (Abuse, Insult, Profanity)	Wiegand, M., Siegel, M. and Ruppenhofer, J., 2018. Overview of the GermEval 2018 Shared Task on the Identification of Offensive Language. In: Proceedings of GermEval 2018, 14th Conference on Natural Language Processing (KONVENS 2018). Vienna, Austria: Research Gate.
Overview of the HASOC track at FIRE 2019: Hate Speech and Offensive Content Identification in Indo-European Languages	German	Twitter and Facebook	4,669	Hate / Offensive or neither, B: Hatespeech, Offensive, or Profane Details of task: Group-directed + Person-directed	Mandl, T., Modha, S., Majumder, P., Patel, D., Dave, M., Mandlia, C. and Patel, A., 2019. Overview of the HASOC track at FIRE 2019. In: Proceedings of the 11th Forum for Information Retrieval Evaluation

Deep Learning for User Comment Moderation, Flagged Comments	Greek	Gazetta, posts	1,450,000	Binary (Flagged, Not)	Pavlopoulos, J., Malakasiotis, P. and Androutsopoulos, I., 2017. Deep Learning for User Comment Moderation. In: Proceedings of the First Workshop on Abusive Language Online. Vancouver, Canada: Association for Computational Linguistics, pp.25-35.
Deep Learning for User Comment Moderation, Moderated Comments	Greek	Gazetta, posts	1,500	Binary (Flagged, Not)	Pavlopoulos, J., Malakasiotis, P. and Androutsopoulos, I., 2017. Deep Learning for User Comment Moderation. In: Proceedings of the First Workshop on Abusive Language Online. Vancouver, Canada: Association for Computational Linguistics, pp.25-35.
Offensive Language Identification in Greek	Greek	Twitter	4,779	Branching structure of tasks: Binary (Offensive, Not), Within Offensive (Target, Not), Within Target (Individual, Group, Other)	Pitenis, Z., Zampieri, M. and Ranasinghe, T., 2020. Offensive Language Identification in Greek. ArXiv.
Aggression-annotated Corpus of Hindi-English Code-mixed Data	Hindi-English	Facebook	18,000	3 part hierachy for hate (None, Covert Aggression, Overt Aggression), 4 part target categorisation (Physical threat, Sexual threat, Identity threat, Non-threatening aggression), 3-part discursive role categorisation (Attack, Defend, Abet)	Kumar, R., Reganti, A., Bhatia, A. and Maheshwari, T., 2018. Aggression-annotated Corpus of Hindi-English Code-mixed Data. ArXiv,.
Aggression-annotated Corpus of Hindi-English Code-mixed Data	Hindi-English	Twitter	21,000	3 part hierachy for hate (None, Covert Aggression, Overt Aggression), 4 part target categorisation (Physical threat, Sexual threat, Identity threat, Non-threatening aggression), 3-part discursive role categorisation (Attack, Defend, Abet)	Kumar, R., Reganti, A., Bhatia, A. and Maheshwari, T., 2018. Aggression-annotated Corpus of Hindi-English Code-mixed Data. ArXiv
Did You Offend Me? Classification of Offensive Tweets in Hinglish Language	Hindi-English	Twitter	3,189	Hierarchy (Not Offensive, Abusive, Hate)	Mathur, P., Sawhney, R., Ayyar, M. and Shah, R., 2018. Did you offend me? Classification of Offensive Tweets in Hinglish Language. In: Proceedings of the 2nd Workshop on Abusive Language Online (ALW2). Brussels, Belgium: Association for Computational Linguistics, pp.138-148.
A Dataset of Hindi-English Code-Mixed Social Media Text for Hate Speech Detection	Hindi-English	Twitter	4,575	Binary (Hate, Not)	Bohra, A., Vijay, D., Singh, V., Sarfaraz Akhtar, S. and Shrivastava, M., 2018. A Dataset of Hindi-English Code-Mixed Social Media Text for Hate Speech Detection. In: Proceedings of the Second Workshop on Computational Modeling of People's Opinions,

					Personality, and Emotions in Social Media. New Orleans, Louisiana: Association for Computational Linguistics, pp.36-41.
Overview of the HASOC track at FIRE 2019: Hate Speech and Offensive Content Identification in Indo-European Languages	Hindi	Twitter and Facebook	5,983	Hate, Offensive or Neither, B: Hatespeech, Offensive, or Profane, C: Targeted or Untargeted	Mandl, T., Modha, S., Majumder, P., Patel, D., Dave, M., Mandlia, C. and Patel, A., 2019. Overview of the HASOC track at FIRE 2019. In: Proceedings of the 11th Forum for Information Retrieval Evaluation
Hate Speech Detection in the Indonesian Language: A Dataset and Preliminary Study	Indonesia n	Twitter	713	Binary (Hate, Not)	Alfina, I., Mulia, R., Fanany, M. and Ekanata, Y., 2017. Hate Speech Detection in the Indonesian Language: A Dataset and Preliminary Study. In: International Conference on Advanced Computer Science and Information Systems. pp.233-238.
Multi-Label Hate Speech and Abusive Language Detection in Indonesian Twitter	Indonesia n	Twitter	13,169	No hate speech, No hate speech but abusive, Hate speech but no abuse, Hate speech and abuse), within hate, category (Religion/creed, Race/ethnicity, Physical/disability, Gender/sexual orientation, Other invective/slander), within hate, strength (Weak, Moderate and Strong)	Okky Ibrohim, M. and Budi, I., 2019. Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter. In: Proceedings of the Third Workshop on Abusive Language Online. Florence, Italy: Association for Computational Linguistics, pp.46-57.
A Dataset and Preliminaries Study for Abusive Language Detection in Indonesian Social Media	Indonesia n	Twitter	2,016	Hierarchical (Not abusive, Abusive but not offensive, Offensive)	Ibrohim, M. and Budi, I., 2018. A Dataset and Preliminaries Study for Abusive Language Detection in Indonesian Social Media. Procedia Computer Science, 135, pp.222-229.
An Italian Twitter Corpus of Hate Speech against Immigrants	Italian	Twitter	1,827	Binary (Immigrants/Roma/Muslims , Not), additional categories. Within Hate, Intensity measurement (Aggressiveness: No, Weak, Strong, Offensiveness: No, Weak, Strong, Irony: No, Yes, Stereotype: No, Yes, Incitement degree: 0-4)	Sanguinetti, M., Poletto, F., Bosco, C., Patti, V. and Stranisci, M., 2018. An Italian Twitter Corpus of Hate Speech against Immigrants. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). Miyazaki, Japan: European Language Resources Association (ELRA).
Overview of the EVALITA 2018 Hate Speech Detection Task	Italian	Facebook	4,000	Binary (Hate, Not), Within hate for Facebook only, strength (No hate, Weak hate, Strong hate) and theme ((1) religion, (2) physical and/or mental handicap, (3) socio-economic status, (4) politics, (5) race, (6) sex and gender, (7) Other)	Bosco, C., Dell'Orletta, F. and Poletto, F., 2018. Overview of the EVALITA 2018 Hate Speech Detection Task. In: EVALITA 2018-Sixth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. CEUR, pp.1-9.



Overview of the EVALITA 2018 Hate Speech Detection Task	Italian	Twitter	4,000	Binary (Hate, Not), Within Hate For Twitter only Intensity (1-4 rating), Aggressiveness (No, Weak, Strong), Offensiveness (No, Weak, Strong), Irony (Yes, No)	Bosco, C., Dell'Orletta, F. and Poletto, F., 2018. Overview of the EVALITA 2018 Hate Speech Detection Task. In: EVALITA 2018-Sixth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. CEUR, pp.1-9.
CONAN - COUNTER NARRATIVES THROUGH NICHE SOURCING: A MULTILINGUAL DATASET OF RESPONSES TO FIGHT ONLINE HATE SPEECH	Italian	Synthetic / Facebook	1,071	Binary (Islamophobic, Not), Multi-topic (Culture, Economics, Crimes, Rape, Terrorism, Women Oppression, History, Other/generic)	Chung, Y., Kuzmenko, E., Tekiroglu, S. and Guerini, M., 2019. CONAN - COUNTER NARRATIVES THROUGH NICHE SOURCING: A MULTILINGUAL DATASET OF RESPONSES TO FIGHT ONLINE HATE SPEECH. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, pp.2819-2829.
Creating a WhatsApp Dataset to Study Pre-teen Cyberbullying	Italian	Synthetic / Whatsapp	14,600	Binary (Cyberbullying, Not)	Sprugnoli, R., Menini, S., Tonelli, S., Oncini, F. and Piras, E., 2018. Creating a WhatsApp Dataset to Study Pre-teen Cyberbullying. In: Proceedings of the 2nd Workshop on Abusive Language Online (ALW2) Month: October. Brussels, Belgium: Association for Computational Linguistics, pp.51-59.
Results of the PolEval 2019 Shared Task 6: First Dataset and Open Shared Task for Automatic Cyberbullying Detection in Polish Twitter	Polish	Twitter	10,041	Harmfulness score (three values), Multilabel from seven phenomena Details of task: Person-directed	Ogrodniczuk, M. and Kobylinski, L., 2019. Results of the PolEval 2019 Shared Task 6: First Dataset and Open Shared Task for Automatic Cyberbullying Detection in Polish Twitter. In: Proceedings of the PolEval 2019 Workshop. Warszawa: Institute of Computer Science, Polish Academy of Sciences.
A Hierarchically-Labeled Portuguese Hate Speech Dataset	Portuguese	Twitter	3,059	Binary (Hate, Not), Multi-level (81 categories, identified inductively; categories have different granularities and content can be assigned to multiple categories at once)	Fortuna, P., Rocha da Silva, J., Soler-Company, J., Warner, L. and Nunes, S., 2019. A Hierarchically-Labeled Portuguese Hate Speech Dataset. In: Proceedings of the Third Workshop on Abusive Language Online. Florence, Italy: Association for Computational Linguistics, pp.94-104.
Offensive Comments in the Brazilian Web: A Dataset and Baseline Results	Portuguese	Twitter	1,250	Binary (Offensive, Not), Target (Xenophobia, homophobia, sexism, racism, cursing, religious intolerance)	de Pelle, R. and Moreira, V., 2017. Offensive Comments in the Brazilian Web: A Dataset and Baseline Results. In: VI Brazilian Workshop on Social Network Analysis and Mining. SBC.
Datasets of Slovene and Croatian Moderated News	Slovene	MMC RTV website, posts	7,600,000	Binary (Deleted, Not)	Ljubešić, N., Erjavec, T. and Fišer, D., 2018. Datasets of Slovene and Croatian Moderated News Comments. In: Proceedings of the 2nd Workshop on Abusive

Comments					Language Online (ALW2). Brussels, Belgium: Association for Computational Linguistics, pp.124-131.
Overview of MEX-A3T at IberEval 2018: Authorship and Aggressiveness Analysis in Mexican Spanish Tweets	Spanish	Twitter	11,000	Binary (Aggressive, Not)	Alvarez-Carmona, M., Guzman-Falcon, E., Montes-y-Gomez, M., Escalante, H., Villasenor-Pineda, L., Reyes-Meza, V. and Rico-Sulayes, A., 2018. Overview of MEX-A3T at IberEval 2018: Authorship and aggressiveness analysis in Mexican Spanish tweets. In: Proceedings of the Third Workshop on Evaluation of Human Language Technologies for Iberian Languages (IberEval 2018).
Overview of the Task on Automatic Misogyny Identification at IberEval 2018	Spanish	Twitter	4,138	Binary (Misogyny, Not), 5 categories (Stereotype, Dominance, Derailing, Sexual harassment, Discredit), Target of misogyny (Active or Passive)	Fersini, E., Rosso, P. and Anzovino, M., 2018. Overview of the Task on Automatic Misogyny Identification at IberEval 2018. In: Proceedings of the Third Workshop on Evaluation of Human Language Technologies for Iberian Languages (IberEval 2018).
hatEval, SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter	Spanish	Twitter	6,600	Branching structure of tasks: Binary (Hate, Not), Within Hate (Group, Individual), Within Hate (Aggressive, Not)	Basile, V., Bosco, C., Fersini, E., Nozza, D., Patti, V., Pardo, F., Rosso, P. and Sanguinetti, M., 2019. SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter. In: Proceedings of the 13th International Workshop on Semantic Evaluation. Minneapolis, Minnesota: Association for Computational Linguistics, pp.54-63.
A Corpus of Turkish Offensive Language on Social Media	Turkish	Twitter	36,232	Branching structure of tasks: Binary (Hate, Not), Within Hate (Group, Individual), Within Hate (Aggressive, Not)	Çöltekin, C., 2020. A Corpus of Turkish Offensive Language on Social Media. In: Proceedings of the 12th International Conference on Language Resources and Evaluation.