Journal on Interactive Systems, 2026, 17:1
ISSN: 2763-7719 • doi: 10.5753/jis.2026.6767

Research Paper

Exploratory Review of Emotion Recognition Resources in Brazilian Portuguese

Neelakshi Joshi [CTI Renato Archer | *njoshi@cti.gov.br*]
Murillo R. Batista [CTI Renato Archer | *mbatista@cti.gov.br*]
Jayant Pendharkar [National Institute for Space Research (INPE) | *jayant.pendharkar@inpe.br*]
Josué J. G. Ramos [CTI Renato Archer | *jgramos@cti.gov.br*]

CTI Renato Archer, Dom Pedro I Highway (SP-65), Km 143,6 - Chácaras Campos dos Amarais, Campinas, SP, 13069-901, Brazil.

Abstract. Emotion recognition (ER) is a preliminary step towards endowing emotional intelligence to machines, which learn it from data. Emotion data resources are pivotal resources for studying and building ER models and an important determinant for their advancement. Brazilian Portuguese (BP) is a dialect of the 8th most spoken language in the world and the 7th globally used on the web, yet it is a low-resource language. Handful of reviews exists citing BP-ER data resources but none are based exclusively on BP nor cover all research literature. This exploratory review assesses the current status of the available BP-ER data resources for advancing ER models in BP. We extensively explored emotion studies among all research publication types up to 2024 and provide an overview of the existing ER data resources in speech, facial expression, and text modalities useful for computation purposes. Overall, 59 data resources were discovered and are listed with the details of emotions included, sources employed in their creation, and their availability under respective modality. The reported data resources are smaller in size and less than 60% are available either openly or on request. A unanimous observation in all the works that created and studied these resources highlights their scarcity. As research with adequate resources will be beneficial for advancing ER modeling, we emphasize the need for the larger in-the-wild multimodal ER data resources in BP. We believe our review will be beneficial in guiding future studies on the aspects of building corpora for advancing emotion recognition modeling in BP, as well as for any low-resource language studies.

Keywords: Brazilian Portuguese, Emotion Recognition Data Resources, Emotion Mining, Review

Edited by: Renan Vinicius Aranha | **Received:** 29 October 2025 • **Accepted:** 15 June 2026 • **Published:** 13 July 2026

1 Introduction

In a societal context, emotions influence cognition, which in turn affects intellect [Cambria *et al*., 2012]. Therefore, modeling emotions is a crucial factor in the development of artificial emotional intelligence (AEI), which aims to confer human-like emotional intelligence to machines, i.e., capabilities to process, interpret, and simulate emotions to enhance human-machine interactions [Schuller and Schuller, 2018; Schuller, 2018; Gao *et al*., 2021]. AEI forms the basis for an affective interactive system that integrates affective computing, human-computer interaction, artificial intelligence, sensing and bio-responsiveness. As a result, it acquires the ability to simulate responses akin to human beings who fine-tune their behavior according to a fellow interlocutor's emotional states during conversations [De Angeli *et al*., 2004; Bucior and Plentz, 2025].

Such interactive systems can be tailored to generate emotionally aware responses with context-adaptive behavior, providing engaged human-machine interaction. Incorporating contemporary technology–large language models (LLMs) and AI agents–can provide personalized services. These systems can be utilized in various ways, including conversational agents, healthcare or therapy assistants, educational tutor agents, customer support, and entertainment [Bucior and Plentz, 2025; Domingues Aparecido *et al*., 2025; De Angeli *et al*., 2004; Ferrada and Camarinha-Matos, 2024; Mendes *et al*., 2024a]. The most critical part of adopting these systems is designing policies to monitor ethics and safety. Already, users prefer bots and applications with AEI capabilities over those without [Vogt *et al*., 2008; Wang *et al*., 2023]. For instance, AI technology adhering to emotional intelligence has been circumstantiated during the recently experienced stressful pandemic period [Sturgill *et al*., 2021; Rossettini *et al*., 2021]. As a result, AEI is of interest both to academia and industry.

Emotion recognition (ER) is one of the building blocks of AEI that identifies human emotions embedded in modalities, viz., audio, visual, text, and physiological signals. These modalities are available as data resources and are essential for ER modeling and backbone to affective interactive system. Over time, ER models improved due to advances in machine learning techniques, processing power, and availability of data resources. Progress in the ER field has resulted in automatic emotion recognition (AER) systems that facilitated wider ER applications across multiple domains, such as education, healthcare, robotics, marketing, advertising, etc. [Saganowski, 2022; Huang and Rust, 2021].

Advances in ER models for real-world applications are driven by the available volume of the data resources and its quality [Schuller, 2018]. It is evident that additional data will enable machines to acquire the intricacies and complexities within it, making them more adept at identifying emotions [Formolo and Bosse, 2017]. However in the context of local population, experiments have reported that the ability to perceive and identify emotions from an unfamiliar language or culture is hindered by the cultural influences on languages,

nonverbal signals, gestures, expressions, and the disparities between cultures [Gomes *et al*., 2024; Cortiz and Boggio, 2022; Silva *et al*., 2016; Tejada *et al*., 2022; Scherer *et al*., 2001]. Thereby pointing to the need of data resources enriched with native linguistic and cultural aspects to enhance the effectiveness of AER systems.

Recent reviews have outlined various aspects of ER modeling, including data resources, state-of-the-art methods, challenges, and future directions [Schuller, 2018; Wang *et al*., 2023; Saganowski, 2022; Elkobaisi *et al*., 2022; Wang *et al*., 2022; Poria *et al*., 2017; Deng and Ren, 2021; Dalvi *et al*., 2021; Spezialetti *et al*., 2020]. These reviews, while providing a global overview of advances in ER, emphasize the need for in-the-wild multimodal emotion data resources in low-resource languages (relative to the English language) to enhance learning for harmonious human-machine interactions.

Portuguese[1] is ranked as the 8th most spoken language in the world and the 7th most used language globally on the internet[2]. Its dialect, Brazilian Portuguese (BP), accounts for 70% of the Portuguese-speaking population. Yet, it is a low resourced language with a handful of reviews citing BP ER data resources — one is based on speech [Elkobaisi *et al*., 2022], other on facial expressions [Fabrício *et al*., 2022] modalities, two based on sentiment analysis [Pereira, 2021; Steiner-Correa *et al*., 2018], and two on hate speech [Fortuna and Nunes, 2018; Poletto *et al*., 2021]. Also, there are two GitHub compilation pages listing two data resources of speech modality[3] and four of hate[4].

These reviews mention the scarcity of data resources in BP but none of them covered all research literature, thus not citing many of the resources nor are based exclusively on BP. Hence, in this exploratory review, we extensively explored affect studies among all research publication types up to 2024 and provide an overview of existing BP-ER data resources in speech, facial expression, and text modalities useful for computation purposes, in unison with affective interactive systems. As earlier reviews [Pereira, 2021; Steiner-Correa *et al*., 2018] focused on sentiment analysis (opinion mining) using text modality, we chose to report data resources that extended their analyses to find underlying emotions from text, also known as emotion mining. Another reason is that we observed numerous ingenious sources used for sentiment analysis, which are worth reporting in a different article.

Besides text emotion mining, work on building resources for hate speech and offensive comments is also included, thereby presenting a holistic view. Given the proliferation of social media usage and contents with prevalent hate (or hatred) and other complex negative emotions, the content moderation task has become a necessary step to filter messages for hate speech or offensive comments. Several research works proposed modeling strategies to classify such contents. As our objective is to find data resources useful for ER modeling, we find it necessary to include studies on hate speech that go beyond sentiments and can categorize. The most widely adopted affective computing model based on cognitive appraisal theory, OCC (see Section 2.2) model [Ortony *et al*., 1988], has classified hate or hatred as a complex emotion, specifically, an object-based emotion arising from negative cognitive evaluation. In literature, several theoretical, experimental, and clinical case studies are discussed on how hate is a complex emotion [Halperin, 2008, 2011; Halperin *et al*., 2012; Marques, 2023]. Furthermore, some works have closely associated anger, contempt, disgust, fear, and anticipation emotions with hatred; and some of these can be emotional antecedents of hate speech [Pascual-Leone *et al*., 2013; Martínez *et al*., 2022; Bilewicz *et al*., 2025; Chavinda *et al*., 2026]. While further discussion on this is outside the scope of this review, we direct readers to the aforementioned literature.

The current work is organized as follows: a brief description of different types of data resources and emotion models that are used in the BP data resources is presented in Section 2. Section 3 describes our search methodology. Section 4 presents the discovered BP ER data resources in different modalities. The findings are discussed in section 5. Finally, section 6 concludes our review.

2 Types of Data Resources and Emotion Models

2.1 Types of Data Resources

ER data resources are categorized based on their data acquisition methods as spontaneous or nonspontaneous [Wang *et al*., 2022]. Spontaneous data resources (in-the-wild or natural) are obtained from real day-to-day conversations. For example, CMU Multimodal Opinion Sentiment and Emotion Intensity (CMU-MOSEI, [Liang *et al*., 2018]) is an in-the-wild dataset that is acquired from YouTube videos. In spontaneous data resources, a way of conveying emotions and its intensity depends on the temperament of the person and context. The availability of natural expressions with the desired emotions and in the desired language is itself challenging. Also, there are legal and ethical issues on building such data resources. Another challenge arises in the case of a non-actor, where the expression of emotion may be vague. Also, emotion expression may differ due to influences of culture, language, ethnicity, and region [Elkobaisi *et al*., 2022; Wang *et al*., 2022; Poria *et al*., 2017; Wierzbicka, 1999; Elfenbein and Ambady, 2002].

Nonspontaneous data resources can be further categorized into acted and elicited data resources. An acted data resources is recorded with an expert actor, who can simulate predefined text or expressions with desired emotions. For instance, Surrey Audio-Visual Expressed Emotion (SAVEE, [Jackson and Haq, 2014]) is an acted dataset. Each data resource is distinct as it varies in the number of utterances, expressions, and number of emotions. It is the most available and may be the most convenient source to obtain all basic or desired emotions in a desired language. Acted data resources are produced in an extensively controlled environment, in-the-lab, regulating an actor's emotional performance. This may result in overplaying the emotion compared to a natural

[1] `https://www.ethnologue.com/insights/ethnologue200/`. Accessed on 06 June 2026.

[2] `https://w3techs.com/technologies/overview/content_language`. Accessed on 06 June 2026.

[3] `https://superkogito.github.io/SER-datasets/`. Accessed on 06 June 2026.

[4] `https://github.com/leondz/hatespeechdata`. Accessed on 06 June 2026.

expression [Wang *et al*., 2022].

An elicited or induced data resources is recorded by inducing desired emotions artificially for a participant on a recording environment. Belfast Induced Natural Emotion [Sneddon *et al*., 2012] is an example of elicited dataset. A trade-off between spontaneity and strict control can be achieved by giving liberty of emotion expression to participants in a controlled environment, in order to obtain emotions naturally [Wang *et al*., 2022]. It is considered to be closer to spontaneous data resource, but depends on how the simulated environment is controlled, otherwise it will not differ from an acted data resource [Schuller, 2018; Vogt *et al*., 2008].

2.2 Emotion Models

Many theoretical frameworks are designed to understand emotions. Here, we succinctly present emotion frameworks to facilitate comprehension of the data resources discussed.

Based on the mode of representation of emotions, emotion models are classified into categorical (discrete) and dimensional (continuous). The former provides a set of distinct emotions. The widely accepted and practiced Ekman's discrete model [Ekman and Friesen, 1971] is based on his controlled experiments where he discovered that the basic six (or big six) emotions — anger, disgust, fear, happiness, sadness, and surprise, are universal across cultures. Other complex emotions can be described using a combination of these big six emotions. However, it fails to account for the interdependence between emotions and languages [Cambria *et al*., 2012].

In the dimensional model, emotions are interrelated across multidimensions in a continuum. One of the most adapted continuous models is Russell's Circumplex model [Russell, 1980], which comprises valence and arousal dimensions. The valence (or pleasure) dimension describes emotion on the pleasant to unpleasant scale. The arousal (or activation) dimension describes intensity or alertness level, scaling from active to passive. Appraisal dimensions (or stimulus evaluation checks) consider the appraisal process that elicited the emotion, thus, associating stimuli to emotional reaction.

Scherer [Scherer, 2005] extended the Circumplex model to the Semantic Space by including appraisal dimensions: goal conduciveness to goal obstructiveness and high to low control/ power scale. The former axis informs the ease in realizing the goal, whereas the latter informs the sense of control or potential. Ortony, Clore and Collins [Ortony *et al*., 1988] proposed the OCC (by initials of authors) model by encompassing logical operations, objects consequences of events, actions of agents, and aspects of objects, implicated in the appraisal process. It presents 22 emotions in a hierarchically organized structure, including valence and emotion intensity parameters. Another continuous emotion model is the hourglass of emotion (HOE) model [Cambria *et al*., 2012] designed to encompass emotions embedded in natural languages. An emotion is scaled on four dimensions: pleasantness, sensitivity, attention, and aptitude, with two different levels of affect categorization for each dimension.

3 Search Methodology

A review of literature exploring available BP ER data resources in the audio, facial expression, and text modalities was conducted using

> Brazilian Digital Library of Theses and Dissertations (BDTD), Scientific Electronic Library Online (SciELO) platform which publishes research in various languages including Portuguese, SBC Open Lib (SOL), Web of Science, IEEE Xplore, Association for Computing Machinery (ACM) Digital Library, and Google Scholar.

Being an interdisciplinary study, we explored literature in various domains such as

> health and biological sciences (e.g., psychology, phonoaudiology, speech therapy, and neurosciences), computational sciences, robotics, linguistics, philosophy, communication and multidisciplinary studies.

And included all research publication types:

> full-text articles in journals, conferences, proceedings, and workshops; theses and dissertations at doctoral, masters, and bachelor levels.

Search terms in English include the following:

> Brazilian Portuguese, emotion recognition database, emotion recognition corpus, recognition of speech emotion, speech emotion recognition, recognition of facial emotion, facial emotion recognition, recognition of text emotion, text emotion recognition, and facebank.

Search terms in BP include the following:

> Português brasileiro, banco de dados, base de dados, banco de faces, banco de expressões emocionais, reconhecimento de emoções, corpus de reconhecimento de emoções, reconhecimento de emoções na fala, reconhecimento de emoções faciais, reconhecimento de emoções de textos.

We observed our search differed when different combination of keywords were used. For example "facial emotion recognition" and "recognition of facial emotion" provided few different results. Most of the literature was found with SciELO, BDTD, SOL, and Google Scholar for keywords in BP. References found in the filtered results as well as their citations were also explored to find relevant articles.
Inclusion criteria:

i data resources pertaining to BP across speech, text or facial expression modalities that are useful for computational processing;
ii published in English or Brazilian Portuguese language and until the year 2024;
iii publication available openly or can access through the Brazilian Education's Ministry research portal (CAPES Periódicos) that gives access to nearly 45000 journals for Brazilian researchers, or when authors provided their article on our request.

Exclusion criteria:

i emotion perception studies that can not be useful for computation purposes are excluded;
ii literature which is behind a paywall and could not be accessed through CAPES Periódicos, and when we did not get the response to our request from those respective authors.

The above defined criteria were employed to search and select the manuscripts for this review, and served as the basis for resolving disagreements, if any, among the authors.

4 Emotion Recognition data resources in Brazilian Portuguese

The included ER data resources useful for computation in three different modalities: speech, facial expressions, and text are described in subsections 4.1, 4.2, and 4.3, respectively. Subsection 4.4 introduces multimodal resources. Also, these search results are listed in Tables 1–4 in each subsection, detailing name of a resource (otherwise subject) along with its year, type, source, emotions and their total number, its status of availability, and reference.

4.1 Speech ER (SER) Resources

The earliest SER corpus we could find was created by Colamarco and Moraes [2008]. It is a small speech act corpus consisting of 16 utterances including statements, yes/no, questions, requests, and orders each expressing the emotions anger, happiness, neutral, and sadness. The recordings were made with a native Brazilian woman and were accompanied by appropriate context. The objective was to study the association between speech acts and emotions based on their respective linguistic and paralinguistic characteristics. Regarding corpus availability, no information was mentioned.

Martins [2011] constructed a corpus for BP speech recognition and synthesis applications, also for emotion recognition. Referring to earlier work, 1000 phonetically balanced phrases were selected from the CETEN-Folha[5] corpus based on BP newspaper contents. From those, two researchers independently selected phrases for SER purposes. A total of 161 phrases per anger, happiness, and sadness emotions were recorded by professional announcers. This corpus was created in three modes: dirty, which contained breathing and leap sounds, clean, where those sounds were removed, and clean with effects that contained clean audio with effects like reverb. No information about the corpus' availability was found, but the used phrases were published in the appendix.

Rosa [2017] built the BP SER dataset[6] by extracting audios from films and YouTube videos of a few seconds duration because the author could not find any existing dataset at that time. It contains a total of 143 audios with anger, fear, happiness, neutral, sadness, surprise, and tense emotions. The author alerted of possible errors in the annotations, as the author himself annotated the extracted audios.

The Voice emotion recognition database (VERBO, [Torres *et al*., 2018]) is an acted SER corpus. It is recorded with 12 professional Brazilian actors (6 females and 6 males) with theater experience of 2 to 3 years and were chosen from different regions in Brazil. It contains 14 predefined phrases, including all BP linguistic phonemes, with four categories: short sentences, phrases, questions, and nonsense phrases. The VERBO comprises a total of seven emotions: the big six and neutral. The clean, semanticless audios are labeled according to the emotions presented in the utterances. It includes a total of 1,167 audio files of 47 minutes duration. Three professional clinical psychologists validated it. The inter-rater agreement with Fleiss' Kappa was found to be 65% and emotions were validated with 76% accuracy. This resource is available on request.

SER corpus Emossônico[7] [Kingeski, 2019] contains three datasets, each obtained from different sources. The first dataset contains 141 audios, which were extracted from 14 national films, 1 national series, and 5 dubbed films composed of 85 female and 56 male voices. Emotions included are anger, fear, happiness, neutral, and sadness. The second dataset contains 206 audios with seven emotions: the big six and neutral, which were selected from online videos of ordinary people (5 males). The third dataset, built by elicited emotions, was created by the author with 6 female and 9 male participants while they were watching videos (a way to induce emotions). A total of 200 audios in seven emotions were recorded: the big six and a neutral state. The process of inducing and collecting emotions was described by the author. The validation process was performed by 20 people. The author obtained the approval from the ethical committee to carry out the experiment.

Lima [2021] recorded the EMOVOX-BR emotion speech corpus to find vocal and speech parameters to distinguish different emotions. The author obtained the approval from the ethical committee to carry out the experiment. The author first collected sociodemographic data of 26 (11 females, 15 males; average age 27 years) native Brazilian professional actors (16) and performing art students (10) using an online form who volunteered for the experiment. Then, three speech tasks were conducted: with sustained vowels; numbers 1 to 10; semi-spontaneous phrases in varied emotions through an online meeting platform using a smartphone. Ten speech therapists evaluated these audios to identify emotions along with their intensities, valence, and auditory characteristics helpful in differentiating emotions. Kappa for inter-judge agreement was found to be 0.711 and all emotions were validated with >70% accuracy. The corpus comprises 182 audios with the best signal-to-noise ratio (SNR >30dB) recorded in the big six and neutral emotions. Loudness and pitch were found to be essential parameters for emotion recognition. No information on its availability was found.

emoEURJ[8] [Bastos *et al*., 2021] was developed at the State University of Rio de Janeiro for the development of SER models acknowledging the scarcity of data resources in BP. Ten sentences were recorded by eight actors (4 females and 4 males) in anger, happiness, sadness, and neutral

[5] `http://www.linguateca.pt/cetenfolha/`. Accessed on 06 June 2026.

[6] `https://github.com/jdarosaj/emotion_portuguese_database`. Accessed on 06 June 2026.

[7] `https://www.udesc.br/cct/geb/projetos/emossonico`. Accessed on 06 June 2026.

[8] `https://zenodo.org/records/5427549`. Accessed on 06 June 2026.

Table 1. Speech Emotion Recognition Resources in Brazilian Portuguese
TYPE: A (acted), S (spontaneous) I (induced); Status: Y (yes), R (upon request), N (No), – (no information)

Resource (year)	Type	Source	Emotions	Number	Status	Ref
Speech acts (2008)	A	Recorded	anger, happiness, neutral, sadness	16	–	Colamarco and Moraes [2008]
SER resource (2011)	A	CETEN-Folha	anger, happiness, sadness	483	–	Martins [2011]
SER resource (2017)	S	Films, YouTube videos	anger, fear, happiness, neutral, sadness, surprise, tense	143	Y	Rosa [2017]
VERBO (2018)	A	Recorded	big six, neutral	1167	R	Torres *et al.* [2018]
Emossônico (2019)	S, I	Films, TV, & Online videos & Recorded	big six, neutral	141, & 206, &200	Y	Kingeski [2019]
EMOVOX-BR (2021)	A	Recorded	big six, neutral	182	–	Lima [2021]
emoEURJ (2021)	A	Recorded	anger, happiness, neutral, sadness	377	Y	Bastos *et al.* [2021]
CORAA (2022)	S	C-ORAL-BRASIL I	neutral, non-neutral-female, non-neutral-male	933	Y	Raso and Mello [2012]

emotions. It comprises 377 audios, where distribution per emotion is as: anger (94), happiness (91), sadness (100), and neutral (92).

CORAA[9] SER version 1.0 is a spontaneous ER corpus of BP utterances collected from the corpus C-ORAL-BRASIL I [Raso and Mello, 2012], which was designed to study spoken BP. This corpus was made available in the SE&R[10] 2022 workshop, an international conference on the computational processing of Portuguese (PROPOR 2022), as a voice emotion recognition competition [Marcacini *et al.*, 2022]. It contains a total of 933 audio files, of which 625 are in the training set and the remaining 308 are in the test set. The total duration of the recorded audio is 60.21 minutes, with a minimum of 2 seconds and a maximum of 14.78 seconds. These audios are labeled into three categories: neutral, non-neutral-female, and non-neutral-male, depending upon the presence of paralinguistic elements in them, and are highly imbalanced. Audio segments without well-defined emotions are categorized as neutral, whereas segments with the presence of one of the primary emotions are categorized as non-neutral. It contains noise and paralinguistic elements like laughter and crying. In addition, based on the training set, two baseline macro F1 scores of 55% using wav2vec features and 50% using prosodic features were provided.

An interesting observation we encountered was that radio programs were used as SER resources; however, no data resource was compiled. Zampar [2010] used radio programs "A Hora do Mução" and "Fala, Gente Fina!" to study how grotesque emotion was used to attract young listeners. Oliveira *et al.* [2014] analyzed the Brazilian radio program "A Hora do Mução" to study how speech rate decreases with anger emotion compared to normal speech rate. Silva *et al.* [2020] analyzed a private customer care dataset but did not mention its source or availability, and referred to it as CustomerPTBR. It consists of audio recordings of customers along with transcripts and includes anger, fear, happiness, neutral, and sadness as emotions.

4.2 Facial ER (FER) Resources

The Brazilian face dataset FEI[11] [Oliveira and Thomaz, 2006; Thomaz and Giraldi, 2010] is the earliest FER resource we found. It contains images with neutral and smiling face expression, varying light illumination, and profile view up to $\pm 180^{\circ}$. It contains 2,800 face images of 200 participants (100 females and 100 males) aged between 19 and 40 years old with distinct appearance, hairstyle, and adornment. The colored images were captured against a white homogeneous background. For each individual, 14 images were captured with profile rotation to the left and to the right. This dataset provides three sets: original images, manually aligned frontal face images with neutral and smiling face expressions, and normalized and equalized cropped frontal face images. The authors believed that the third set can be useful for experiments with realistic expression synthesization and age estimation. Construction of the dataset and the alignment of images are described in [Oliveira and Thomaz, 2006].

Guimarães [2014] built a dynamic FER resource in the form of micro-videos for 7 basic emotions. To record an experiment a display screen, chair, and camera were kept at a fixed position focusing on the participant' face. A projection screen and a retractable visograph of matte white vinyl fabric was used in the background for standardization. For elicitation purposes, thirteen videos showing six basic emotions (happiness, anger, sadness, fear, surprise, and disgust) with varying intensity and time varying from 1 to 6 minutes

[9] `https://github.com/rmarcacini/ser-coraa-pt-br/`. Accessed on 06 June 2026.

[10] `https://sites.google.com/view/ser2022/home`

[11] `https://fei.edu.br/~cet/facedatabase.html`. Accessed on 06 June 2026.

Table 2. Facial Expression Recognition Resources in Brazilian Portuguese
TYPE: A (acted), S (spontaneous) I (induced); MODE: A (audio), IMG (image), S (sensors), V (video); Status: Y (yes), R (upon request), N (no), – (no information)

Resource (year)	Type	Source	Emotions	Mode (Number)	Status	Ref
FEI (2006)	A	recorded	neutral and smiling face profile images ±180°; varying light illumination	Img (2800)	Y	Oliveira and Thomaz [2006]
FE (2014)	S, A	recorded, audio-visual stimuli	big six	V with 3 intensity levels	–	Guimarães [2014]
CEPS (2015)	A, I	recorded, audio-visual stimuli	big six with 3 intensity levels, and neutral	Img (225)	–	Romani-Sponchiado *et al*. [2015]
Baby FE (2016)	S	recorded	disgust, happiness, pain, sadness, neutral	Img (65)	–	Nascimento [2016]
Game player FE (2017)	S	recorded, video games	big six, neutral, fun, frustration, immersion	V	R	Vieira [2017]
YEPS (2018)	A, I	recorded, audio-visual stimuli	big six, neutral	Img (42)	–	Novello *et al*. [2018]
Game player FE (2018)	S	recorded, video game	positive, negative, neutral	I, V, S	–	Siqueira *et al*. [2018]
Children IRTI (2019)	I	recorded, audio-visual stimuli	disgust, fear, happiness, sadness, surprise	V, Img	Y	[Goulart *et al*., 2019a]
Children IRTI (2019)	I	recorded, social robot	disgust, fear, happiness, sadness, surprise	V, Img	–	[Goulart *et al*., 2019b]
Baby Faces (2019)	S	recorded	anger, fear, happiness, sadness, surprise, neutral	Img (72)	R	Donadon *et al*. [2019]
FABS (2019)	A, I	recorded, audio-visual stimuli	big six, contempt, neutral	Img (1014), V (971)	R	Negrão [2019]
Political propaganda (2020)	I	recorded	big six, contempt, neutral; valence and arousal	Img	–	Dias [2020]
MIGMA (2021)	A, I	recorded, emoticons	big six, contempt, neutral	Img (1,500)	N	Matias *et al*. [2021]
Transcultural FE (2022)	A	recorded, computer assisted tasks	big six, debauchery (or mockery), neutral	Img (806)	R	Tejada *et al*. [2022]
Game player FE (2023)	S	recorded, video game	anger, calm, happiness, sadness	I, V, S	N	Siqueira *et al*. [2023]
Face bank (2023)	A, I	recorded, audio-visual stimuli	big six, neutral	Img (56)	–	Fabrício [2023]
Face bank (2024)	A, I	recorded, audio-visual stimuli	big six, neutral	Img (56)	R	Morais *et al*. [2024]
FBioT (2024)	S	YouTube	happy and neutral	V (70)	Y	Souza *et al*. [2024]

were selected from YouTube. The intense emotional videos were used for men considering literature finding that women are more expressive compared to men. The links of these videos were provided in the work. Along with these videos, images of varying six basic emotions with three different intensities (20%, 60%, 80%) were used as stimuli. The author obtained the approval from the ethical committee to carry out the experiment.

For the experiment, twenty-three volunteers (56.53% women) of age 19-35 (24.04±4.28) years were employees and students of Universidade Federal do Vale do São Francisco (UNIVASF). They were advised not to keep beard or use glasses, hats or bangs that covered the eyes. In the starting researcher explained the procedure and informed them a false objective as a focused attention to avoid any forced/exaggerated expressions. The participants were advised not to touch face, speak or take eyes off the screen during recording. The spontaneous expression videos were recorded using obtained elicitation videos. With the help of two volunteers, videos of acted expressions with varying intensities (weak, moderate, and strong) were recorded.

Later, while showing their recording to participants, the researchers informed the true objective. With the granted consents the recordings were retained otherwise deleted in their presence. Then researchers created 4 sec micro-videos for each emotion per participant from the recorded videos of spontaneous and acted expressions. These videos were evaluated for desired expressions and their genuinity by four judges who were experts in emotion expression patterns. Videos with at least 75% agreements were accepted. Next, another three judges evaluated these videos for emotion expression and their intensities. The videos with at least two judges in-agreement were selected for the final step. Various analyses like psychometric, item response theory, similarity structure analyses were carried out with the help of 201 university students to validate this created resource. The author believed that along with FER studies, this resource will be useful for therapeutic use but do not mention about its availability.

Romani-Sponchiado *et al*. [2015] developed the child emotions picture set (CEPS) with 17 school children in the age groups: 6–7 years, 8–9 years, and 10–11 years, as children can converse within the same age group with ease. As per facial features, participants were from indigenous, Afro-American, and Caucasian ethnicity. Participants were non-actors from southern Brazil. The CEPS covers 6 basic emotions with 3 levels of intensities: low, medium, high, plus neutral in a posed and spontaneous mode. The authors obtained the approval from the ethical committee to carry out the experiment.

For spontaneous images, short videos selected by external child health professionals were used to elicit happiness, fear, disgust, surprise, and neutral emotions. To capture elicit anger and sadness, participants were asked to recollect relevant experiences. For posed images, participants were asked to reproduce the same emotion they felt when exposed to stimuli with 3 different intensities. In preprocessing, all images were checked for eye focus, sharpness, and contrast; converted to black and white; and resized to 300×300 pixels (100 dpi). These images were annotated for emotion and intensity by 30 psychologists with expertise in child development, using online survey software with a forced-choice method. Among the judges, an overall 85.06% agreement was reached for the final dataset. It consists of 225 images of 8 boys and 9 girls where 40% (90) images are spontaneous and 60% (135) are posed. As per facial features of participants, 146 images are of Caucasian children, and 79 are of different ethnic groups: Afro-American and indigenous. No information is given on data availability.

Face expressions of infants between 4 and 24 months were recorded by Nascimento [2016] to capture disgust, happiness, pain, sadness, and neutral emotions. The author obtained the approval from the ethical committee to carry out the experiment. Intramuscular injection vaccine was used as an inducing method to evoke pain. Other emotions were induced with the help of the mother of the infants using toys, gestures, the mother's vocalization, and introducing acidic (lemon) tests. The recording took place in a quiet environment, the vaccination room, and the position and angle of the camera were adjusted as per the position of the baby. Four pediatrics judged images for emotions and quality. Agreement among the judges in the measure of Kappa was not high. Untrained multiparous and primiparous mothers validated images, and the mean emotion validation rate was found to be 68.22%. After applying the selection criteria, a total of 65 images of 13 babies (7 boys and 6 girls) that received a higher recognition rate from evaluators were selected for the dataset. The author believes that this dataset can be a useful resource for health professionals, but no details are provided on how to acquire it.

The objective of Vieira [2017] was to determine whether visual cues – facial expressions, camera proximity, and blinking rates, obtained with a standard webcam while volunteers were playing video game, can effectively measure a player's level of enjoyment (or fun), frustration, immersion and other emotions, without relying on additional sensor data or game metrics. The author chose three open source and unbiased games of different genre among adventure, comedy, platformer, antiquity, family, puzzle, and horror, which were popular and independent of language, if that differs from Portuguese. The author obtained the approval from the ethical committee to carry out the experiment.

The experiment was conducted with total of 35 (17 males and 18 females) adult volunteers (students or employee), who did not have a history of photosensitive epilepsy or repetitive strain injury and were comfortable to carry out experiments, were recruited from a private music school and the campus of the University of São Paulo. Among them eight participants did not grant authorization to reproduce their face images in any media. Among all volunteers 17 wore glasses and 9 male participants had facial hairs. Their average age was 23 for males and 32 for females.

Each game session was of 20 min, 10 min of game play and next 10 min for self-game-play review and answering questionnaires those assessed the levels of (i) frustration, immersion, and fun; (ii) competence, sensory and imaginative immersion, flow, tension/annoyance, challenge, negative and positive affects while playing the game. After playing all sessions, the author collected information on their playing habits and whether these games had been played already. Experi-

ment was conducted in an artificially lit and well aired private room in the music school. The authors created a tool to conduct an experiment and save the data without human intervention. Video was recorded in MPEG-4 compression format with standard 1280 x 720 (HD) resolution and 30 fps.

The author considered two features, immersion and blink rate, to assess player' experience. When player is immersed in a game, may lean towards the game screen and blink rate (blinks per minute) reduces. Immersion is measured as the gradient of the estimated distance between the face and the camera. Blink rate was calculated with the help of extracted facial landmarks and then finding vertical movements of eyelids. FER model trained on two different data resources was used to identify 7 basic emotions using facial landmarks extracted from sequences of the recorded video frames. Then with identified anger, fear, and sadness emotions the level of frustration was assessed. Finally, using all emotions, frustration and immersion levels, the level of fun was assessed. Availability of this data resource is conditional.

Novello *et al*. [2018] developed a youth facial expression dataset, youth emotion picture set (YEPS), including all range of teenagerś in the age group 12–20 years. This includes posed and spontaneous facial expressions for six basic emotions and neutral, consisting of 42 images, of boys and girls from different racial backgrounds. The author obtained the approval from the ethical committee to carry out the experiment. For image acquisition, participants wore a black cape to hide their clothes, and were allowed to wear only small earrings.

The authors used three different methodologies to record their dataset: emotive scenarios (ES), where emotional audios in random order were used for elicitation; posed expressions (PE), where images of basic emotions with varying intensities (15%, 50%, and 100%) were shown; and reaction to visual stimuli (RS), where a difficult level flash game with cheap tricks was used to elicit anger and short movies to induce other remaining emotions. To finalize these stimuli, the authors conducted a pilot study with 10 participants of both gender and age range. Three FER expert psychologists filtered images for quality and emotions and a final set of 42 images were selected to satisfy gender, emotion intensity, and racial background representing Brazilian society. Further, these images were preprocessed, converted to black and white, and resized to 369×475 pixels. Four FER expert psychologists judged the images for emotions which were presented in random order on online platform and high agreement among them was observed, with $\kappa = 0.972$. No information is made available on accessing it.

Goulart *et al*. [2019a] constructed a dataset of facial thermal images of 28 children (16 boys and 12 girls) from Vitoria, Brazil in the age group 7–11 using Infrared Thermal Imaging (IRTI) by measuring changes in facial emissivity for five emotions — disgust, fear, happiness, sadness, and surprise on both valence and arousal dimensions. The author obtained the approval from the ethical committee to carry out the experiment. Under supervision of a psychologist, the authors obtained five videos of audiovisual stimuli per emotion of duration varying from 40 to 130 seconds. Plus, one more positive emotional video was selected for training. The experiment was performed individually with the children in the morning session, with room temperature maintained at 22°C. The children were examined for their mental and physical health and body temperature to discard any influential factors that could alter the thermal analysis while acquiring baseline recordings. A 19-inch screen and a thermal camera were placed at a distance of 85 cm. A Therm-App infrared thermal imaging camera of 8.7 Hz frame rate was used and images were acquired with a spatial resolution of 384×288 ppi with pixel intensity varied from 0 to 255 where darker pixels showed lower emissivity and vice versa.

Initially, positive emotional videos selected for training were shown to the children followed by displaying the stimuli for disgust, fear, happiness, sadness, and surprise emotions respectively interspersed with a blank display for 4s. After displaying each stimuli, children informed perceived emotion on a scale of 1 to 9 relative to valence (pleasure) and arousal (intensity) dimensions. A 19-inch screen and a thermal camera were placed at a distance of 85 cm. A Therm-App infrared thermal imaging camera of 8.7 Hz frame rate was used and images were acquired with spatial resolution of 384×288 ppi with pixel intensity varied from 0 to 255 where darker pixel showed lower emissivity and vice versa, and converted to black and white (BW) images to facilitate detection of facial regions of interest (ROIs). These ROIs had shown emissivity variation per emotion, confirming the desired effect of audiovisual stimuli among the children. The authors described their classification analysis and provided baseline results with mean accuracy >85% and κ >81% for five emotions. The used audiovisual stimuli[12], dataset[13], emotion classification codes are provided by the authors and their links are provided in the article.

Later Goulart *et al*. [2019b] performed a similar experiment proposing a new methodology to overcome the poor resolution of IRTI camera. The author obtained the approval from the ethical committee to carry out the experiment. In this experiment, they used a mobile social robot, N-MARIA (New-Mobile Autonomous Robot for Interaction with Autistics), placed at a distance of 70 cm, as an emotional stimulus during a robot-children interaction and recorded children' facial expressions with two cameras attached to it, a visual (RGB) and an infrared thermal imaging (IRTI) with 2 fps sampling rate to acquire five emotions - disgust, fear, happiness, sadness and surprise. The experiment was carried out in a school (Vitoria, Brazil) with 17 children (9 boys, 8 girls) in the age group of 8 to 12 years old. The room where the experiment was conducted had constant luminous intensity and temperature was kept between 20° to 24° C. The baseline recording was acquired with the robot face covered, followed by a 3-minute interaction. At the end of the experiment, children responded to a questionnaire expressing their emotions and opinions about the robot. No statement regarding availability was made.

Facial expression dataset Baby Faces [Donadon *et al*., 2019], contains 72 colored images of 12 infants capturing the basic six expressions. The author obtained the approval

[12] `https://doi.org/10.1371/journal.pone.0212928.s001`. Accessed on 06 June 2026.

[13] `https://doi.org/10.1371/journal.pone.0212928.s002`. Accessed on 06 June 2026.

from the ethical committee to carry out the experiment. The infants (6 boys and 6 girls) in the age group 6–12 months belonged to Caucasian (66%), black (17%) and Japanese (17%) ethnicity. Elicitation methods were as follows: i) happiness: playing with baby with his/her favorite toy; ii) anger: take away the toy from baby with which it was playing; iii) sadness: parents left the room, leaving baby alone in a room for some time; iv) fear: approach of a stranger to the baby in the absence of respective parents; v) surprise: an audible noise of horn up to 80 dB where parents are out of sight of the baby. vi) neutral: keeping baby in resting condition where parents looked towards baby with neutral look. With the camera being 1 meter away, videos were collected in-the-lab under a controlled environment (temperature, light, and noise). From each video, respective stimulus frames were selected. From those frames, the best frames per emotion per infant were chosen by the authors in agreement, and processed for standardization in terms of size, orientation, and background. The dataset was validated with 119 adult volunteers varying on age, education level, gender, and race. The participants' samples were chosen to represent the variation in Brazilian society. It is available on request.

Another early childhood dataset, namely, the Facial Affective Brazilian Set—Children image and video database (FABS, [Negrão, 2019]) was developed with children of age 4 to 6 years. The author obtained the approval from the ethical committee to carry out the experiment. The participant children were chosen from an acting agency and were Caucasian, African, or Asian descent. Elicitation methods, attention-getting, age-appropriate cartoon excerpts and facial expression photos with desired emotions along with guided imagination (phrases) were selected by consulting professionals and conducted a pilot study to finalize those. For recording, children were without make-up and wore white t-shirts. A white panel was used in the background and the camera was kept at a distance of 60 cm. With the help of stimuli, videos were recorded for big-six, contempt, and neutral. One researcher identified the emotion expression moments for videos and images. An expert in facial analysis evaluated the selected videos and images and with 100% in agreement were retained in FABS. Another two experts in facial analysis validated the FABS. Kappa index for inter-rater agreement was found to be 0.70 and with 73% accuracy.

The FABS contains 1,985 stimuli of 124 children (among 55% girls), of which 1014 (51%) are photographs and 971 (49%) are videos. Among children, ethnicity was found as 71% Caucasian, 24% Afro-descendant and 5% Asian. Emotion distribution in FABS is as follows: anger (234), contempt (276), disgust (310), fear (269), happiness (437), neutral (150), sadness (183), and surprise (126). The author believes that this resource may help characterizing several neurodevelopmental and psychiatric disorders among children in early childhood. Later, its subset was published under the name ChildEFES [Negrão *et al*., 2021] where 100% agreement was reached among evaluating judges. The ChildEFES contains 1,381 stimuli of 124 children. It consists of the following emotions with a respective number of stimuli: anger (157), contempt (183), disgust (170), fear (152), happiness (363), neutral (87), sadness (144), and surprise (104). It is available on request.

Dias [2020] explored the influence of visual elements of election advertisements by creating political propaganda videos with different scenarios in a controlled environment (lab) and analyzing voters' emotional responses using their facial expressions and opinions. The videos were recorded with the help of an actor and randomly included four different digitally inserted backgrounds and words related to political agendas. These videos were used as an emotion elicitation method. Participants' facial expressions along with their intensities were categorized into the big-six, contempt, and neutral, as well as valence and arousal. Performing statistical analysis, the author demonstrated that the visual elements of election advertisements influence the voters emotionally. No information on dataset' availability was provided.

Matias *et al*. [2021] created the MIGMA dataset of 1.5k high quality spatial resolution images with induced and non-induced emotions: the big-six, contempt, and neutral. The author obtained the approval from the ethical committee to carry out the experiment. University staff and students, a total of 323, of varied ethnicity, age, and gender, volunteered for this experiment. Among them, 200 were men, 119 were women, and 3 did not mention gender, and were in the age group 18–55 years. Ethnicity wise distribution was 86.06% white, 2.47% mulatto, 1.23% Asian, 0.3% mixed colors, 2.78% did not declare. The authors collected demographic information and psychiatric symptoms from the participants using a web based system. Under supervision of a psychiatrist, participants were examined for their psychiatric symptoms. Image recording sessions were carried out under a non-induced and induced environment where the participants were asked to capture their photo while imagining emotions, then with the help of emoticons, respectively. The experiment was carried out in the controlled environment by standardizing the background and camera. Images were checked for any inconsistency and were recorded again when necessary, and preprocessed for eye alignment and regularity of the expression by considering mean faces per emotion. They provided a baseline accuracy of 44% for the validation set (15%) where CNN model was trained on 70% for the training set. This resource is not available as informed by the authors.

The first transcultural facial expression dataset for the big six, debauchery (or mockery), and neutral (or at rest) emotions, was constructed by Tejada *et al*. [2022]. This was developed with Brazilian (60) and Colombian (50) non-actor volunteers while they were performing computer-assisted tasks. The average age of participants was 22 years and among them 54 were women. The participants followed the instructions displayed on the screen for live recording by a webcam with 30 FPS rate. The videos were preprocessed to obtain grayscale 350×350 pixels images and their 68 facial landmarks. The authors built this dataset of 806 images very carefully, considering all possible challenges in the process. The final images of facial expressions were chosen, applying many criteria like correct facial emotion expression as participants being non-actors, the corresponding facial landmark identification per emotion and their cross-validation with trained classifiers before they were judged by experts. Five professors each, from Brazil and Colombia, experts in human sciences, evaluated this dataset using the online sur-

vey application provided by the authors.

The authors' motivations in building their resource are as follows: (i) No available dataset includes Brazilian culture. The majority of them are from the USA or Europe, which do not account for pan-cultural behavioral differences. As a result, ER models trained on such data neither provide expected results nor perform well with other cultures. (ii) The authors emphasized the need for a new dataset because the facial expressions during social interactions and responses to computer-assisted tasks are distinct. The authors have detailed and shared their cost-effective resource construction and validation procedures, along with the scripts[14]. They have also encouraged other communities to build their own datasets to study cultural behavior, which will further help in improvements of pan-trans-culture ER models. This dataset is available on request.

Siqueira *et al*. [2018, 2023] conducted experiments with a racing game to explore the impact of game settings on player's ability and assess the player' experience. The authors chose a combination of two configuration parameters in a game, drift (settings: enable or disable) and gearbox (settings: manual or automatic), providing two difficulty levels to study its effect on players' experience. These configurations were introduced randomly, and each participant played all four configurations. They used a webcam to capture facial expressions and a sensor to acquire biosignals during gameplay.

Volunteers were selected through social networks including students to game developers in the age group 18–32 years for these experiments. Each participant had at least some experience of playing a racing game. Twenty (4 females and 16 males) and thirty (10 females, 20 males) candidates participated respectively in the first [Siqueira *et al*., 2018] and second experiment, [Siqueira *et al*., 2023] which were carried out in the department of Computer Science of University of Brasília. All participants were informed about the play test, purpose of experiment, data collection and handling procedure, and their rights. Later they were asked to sign consent form and filled the questionnaire. The authors used facial landmarks as features and annotated basic emotions and neutral emotions using a developed model. Fear and surprise emotions were excluded given low data density.

The authors [Siqueira *et al*., 2018] grouped anger, disgust, sadness as negative, happiness as positive, and neutral as neutral emotions. Further modeled on valence-arousal axes and analyzed using Fuzzy logic. They noted that when game setting matched the capability of player, the player felt satisfied (calm/positive emotion). Thus, they demonstrated possible way to estimate the player's experience. No information is given on data resource availability.

The authors [Siqueira *et al*., 2023] created a facial expression annotation tool with a list of 32 emotions and 10 game events. Two expert social psychologists and sixteen psychology students evaluated facial expressions and respective game events by watching game session recordings in two levels. The first, using all emotions from the given list and the latter, using eight emotions: anger, anxiety, calm, desire, disgust, fear, happiness, and sadness. The game sessions with Fleiss' kappa coefficient >0.6 for inter-rater agreement were used to finalize a dataset. With 8 emotions, 69.5% average agreement was obtained. The final dataset contained data for anger (100), calm (180), happiness (210), and sadness (170) only. Other emotions were discarded due to very low instances. Then, the same procedure was performed with 20 more participants. Their self-reported emotions and self-evaluation using the annotation tool were used to validate the authors' ANN model, which were trained with earlier data to classify emotions. The authors reported 64% mean accuracy as a baseline result. Data access is restricted, but they made their code[15] available.

For FER dataset (face bank) creation, Fabrício [2023] conducted a pilot study to select elicitation methods with images and scenarios for the big six and neutral emotions, and to standardize the database capturing procedure where white photographic panel was used as a background while all the participants used black t-shirt. Using selected stimuli, videos were recorded with experienced actors and non-actor volunteers and preprocessed for quality. It included frame alignment with nose tip (at the center), eye, head positions; adjustment in color, contrast, temperature, shadows, etc.; removed piercings and teeth braces, and resized to 270×360 pixels. The author obtained the approval from the ethical committee to carry out this experiment.

Five expert judges in emotions and psychometry annotated the frames with emotions and annotation difficulty level as easy, medium, and hard. All images with hard annotation level were removed as it would be hard and confusing for ER tasks. Images with easy level annotation with 80% or more inter-rater agreement were chosen for face bank. To comply with the demographic distribution in Brazilian society, images with higher inter-rater agreement were selected from medium level. The dataset contains total of 56 images, 8 images per emotion, of 29 volunteers (16 actors, 13 non-actors) with following demographic distribution in gender (58.6% females and 48.3% males), ethnicity (48.3% white, 34.5% brown, 10.4% black, 3.4% yellow, 3.4% indigenous), and age (51.7% of 18–39 years, 34.5% of 40–59 years, and 13.8% of 60 or more years). The dataset is published in Appendix VI of the thesis [Fabrício, 2023], but without labels and no mention on how to acquire it.

Morais *et al*. [2024] created a Brazilian FER dataset of 56 high quality color images accounting proportion of age, sex, and race in the Brazilian population with big six and neutral emotions. They conducted a pilot study with six mental health professionals and psychology researchers to finalize stimuli images and hypothetical situations for elicitation purposes. The author obtained the approval from the ethical committee to carry out the experiment. It was created with 29 participants (16 actors and 13 non-actors). For the capturing process, they chose white background with white light where all participants wore a black t-shirt, without makeup and any distracting things. Using created online forms, five experts in emotions and psychometrics (3 females, 2 males with master and doctoral degree) validated images to identify emo-

[14] `https://github.com/julian-tejada/EmotionRecognition`. Accessed on 06 June 2026.

[15] `https://github.com/eltonsarmanho/PGD-Ex`. Accessed on 06 June 2026.

tions and its identifying level in easy, medium, or difficult categories. A final dataset of 56 images was constructed: all difficulty category images were discarded; images with 80%-100% inter-rater agreement for easy category were considered; to obtain emotions with eight faces per emotion and to satisfy criteria images were chosen with majority vote from medium category. These images were preprocessed with professional software to standardize positions, lightning, size (270×360 pixels), and to remove piercings and dental applications. The dataset with low resolutions and their descriptive analyses were provided in the supplementary material and the dataset is available on request.

Souza *et al*. [2024] constructed a multivariate Facial Biosignals Time–Series, FBioT[16], dataset using 42 online and 28 downloaded videos from YouTube. The FBioT is a spatiotemporal normalized and stabilized dataset where facial landmarks were correlated to facial action coding system (FACS) to identify the underlying neutral and happy emotions. They developed visual–temporal facial expression recognition (VT-FER) methods that assist automated dataset construction processes with online or offline videos. Steps (i)-(iv) were followed, where the output of an earlier step was fed to the next: (i) metadata extraction; (ii) 68 facial landmarks extraction using Dlib; (iii) creation of stabilized and normalized temporal series; (iv) correlating the temporal series with facial action unit systems. Following the above steps, 22 standardized measures were obtained for seed videos of emotion expression, which were used for annotation.

As this dataset is based on mapping of FACS to facial landmarks, the authors assured that their dataset contains natural face diversity without any bias like race, color, gender, etc. Also, it avoids technical constraints such as lighting and noise conditions, need of frontal position of face, minimal movement restrictions, which traditional datasets experience while capturing temporal-spatial descriptors concurrently. Creating time series instead of images/videos consume less computational resources. The authors provided their baseline binary classification F1-score of 80% using LSTM neural network and compared their result with other two video-FER datasets. They made their codes and detailed documentation[17] available along with the dataset in their GitHub repository.

4.3 Text ER (TER or Emotion Mining) Data Resources

A lexicon, or an emotion dictionary, is an important linguistic tool for emotion and opinion mining, which maps akin words to emotions and opinions, respectively. As the majority of the resources are available in the English language, the following lexicons were developed especially for Brazilian Portuguese. Pasqualotti and Vieira [2008] built the first Brazilian Portuguese affect lexicon WordnetAffectBR, in three stages composed of 289 words. The OCC emotion model relates emotions to cognitive processes. Thus, it was used as a conceptual base and lexical reference to identify words having meaning in the emotion domain. Further, this word base was extended using the following lexical resources: Wordnet, Base Affect, and Wordnet Affect. Wordnet is an English lexical database composed of words and sets of synonyms (synsets). It is organized as per their form and meaning, encompassing noun, verb, adverb, and adjective categories. The base Affect and Wordnet Affect were used to relate the lexical and semantic content of words in Wordnet to affective meaning based on their characteristics and contents, resulting in extending the emotion word database. Next, it was carefully translated to Brazilian Portuguese based on the significance of the words and structure of the synsets where each word was cross-referenced at each stage with the help of experienced translators.

A domain-independent lexicon, OpLexicon[18], was developed by Souza *et al*. [2011], integrating three lexicons obtained with three different methods–a Brazilian Portuguese corpus of movie reviews and PLN-Br CATEG corpus []; a TEP thesaurus [Maziero *et al*., 2008] based on synonymy and antonymy; and translating an English opinion lexicon [Hu and Liu, 2004] to Brazilian Portuguese. It contains 7,077 polarity words, among adjectives, verbs, nouns, and expressions categories. LIWC-PT (or LIWC_2007pt as referred in Carvalho *et al*. [2019]) lexicon [Balage Filho *et al*., 2013] was created by translating the original Linguistic Inquiry and Word Count (LIWC) lexicon from an English language [Pennebaker *et al*., 2001]. A total of 127,149 instances are ascribed to 64 categories for emotion and opinion mining. To improve this lexicon, to add several new categories, and to add words matching linguistic, social, and psychological characteristics, the LIWC_2015pt[19] lexicon was created by translating an English 2015 version of the LIWC lexicon [Carvalho *et al*., 2019, 2024]. It composed of total 73 categories. Based on an English 2015 version of LIWC lexicon, the AffectPT-br[20] lexicon was developed focusing on affect categories [Carvalho *et al*., 2018] using an online translator and an English-Brazilian Portuguese bilingual dictionary. It consists of a total of 1139 words assigned in the affect category. All the creators of the above-mentioned lexicons validated and evaluated it by comparing it against existing lexicons.

Martinazzo [Martinazzo, 2010; Martinazzo and Paraiso, 2010; Martinazzo *et al*., 2011] manually constructed an emotion word dictionary in BP. For this work, to use formal language, a total of 1,002 short news along with their headlines were extracted from a news website. Words were classified into the big six and neutral emotions. A group of 13 volunteers (researchers, doctoral and undergraduate students) annotated and evaluated the predominant emotion in news. A news corpus is available under the Emoções.BR[21] project on request. A Kaggle user presented this emotion dictionary

[16] `https://github.com/jomas007/biosignals-time-series-dataset/`. Accessed on 06 June 2026.

[17] `https://github.com/jomas007/biosignals-time-series-dataset/wiki/Blocks-Description`. Accessed on 06 June 2026.

[18] `https://www.inf.pucrs.br/linatural/wordpress/recursos-e-ferramentas/oplexicon/`. Accessed on 06 June 2026.

[19] `https://github.com/LaCAfe/LIWC2015pt`. Accessed on 06 June 2026.

[20] `https://github.com/LaCAfe/AffectPT-br`. Accessed on 06 June 2026.

[21] `https://www.ppgia.pucpr.br/~paraiso/mineracaodeemocoes/`. Accessed on 06 June 2026.

Table 3. Text Emotion Recognition Resources in Brazilian Portuguese
TYPE: A (acted), S (spontaneous) I (induced); Status: Y (yes), R (upon request), N (no), – (no information)

Resource (year)	Source	Emotions/Affect	Number	Status	Ref
News (2010)	News	big six and neutral	1,002	R	Martinazzo [2010]
News (2013)	News	big six and neutral with low, medium, and high intensities	2,000	R	Dosciatti *et al.* [2013]
Kindle reviews (2014)	Kindle	big six, anticipation and trust	150	–	Santos *et al.* [2014b]
Parallel corpus (2014)	Music lyrics	big six, anticipation, and trust with intensities	840 stanzas	–	Santos *et al.* [2014a]
OffComBR-2 & OffComBR-3 (2017)	News Comments	racism, sexism, homophobia, xenophobia, religious intolerance, and cursing offenses	1,250 & 1,033	Y	De Pelle and Moreira [2017]
MQDEmotion (2018)	Meu Querido Diário	big six	79,523	Y	Nascimento *et al.* [2018]
HateSpeech (2019)	55chan, Twitter	offensive (anger and anxiety), non-offensive (positive emotions)	7,672	Y	Nascimento *et al.* [2019]
Vignettes (2019)	Author' creation	anger, disgust, fear, happiness, gratitude, guilt, sadness, and neutral	40	Y	Wingenbach *et al.* [2019]
hierarchical multiple label hate speech (2019)	Twitter	hierarchical multiple label of fine-grained 81 categories of hate speech	5,668	Y	Fortuna *et al.* [2019]
stock market domain (2020)	Twitter	joy, sadness, anger, fear, trust, disgust, surprise, anticipation	4,553	Y	da Silva [2020]
ToLD-Br (2020)	Twitter	Non-toxic, LGBTQ+phobia, obscene, insult, racism, misogyny, & xenophobia	21k	Y	Leite *et al.* [2020]
GoEmotion-PT-BR (2021)	Twitter	28 emotions	47,405	Y	Cortiz *et al.* [2021]
GoEmotion-PT-BR (2021)	GoEmotion-EN	28 emotions	54,263	Y	Hammes and de Freitas [2021]
MOOC posts (2022)	MOOC	anger, dislike, fear, happiness, sadness, and surprise	766	–	Souza [2022]
COVID-19 vaccine related (2022)	Twitter	indignation, enthusiasm, happiness, pride, suspicion, cynicism, defiance, hope, anger, envy, empathy, disgust, and sadness	85	–	Penteado *et al.* [2022]
HateBR (2022)	Instagram	9 offensive categories in 3 intensities	7,000 & 727	–	Vargas *et al.* [2022]
soccer team related (2023)	tweets	support, anger, happiness, and funny	–	Y	Rodrigues *et al.* [2023]
city related (2023)	tweets	aversion, happiness, sadness, surprise, disgust, desire, none; categories:propaganda, information	4183	–	de Oliveira Moraes [2023]
Election news (2023)	Twitter	Sentiment and big six emotions	53,932	R	Silva *et al.* [2023]
COVID-19 vaccination (2023)	Facebook	35 emotions	523	–	Fernandes-de Oliveira *et al.* [2023a]
COVID-19 vaccination (2023)	Instagram	30 emotions	620	–	Fernandes-de Oliveira *et al.* [2023c]
Hate speech database & lexicon HALEX (2023)	Twitter	Hate, non-hate	33,771	Y	Weitzel *et al.* [2023]
TuPy-E (2023)	Twitter, Instagram	aggressive in: not hate, hate & hate in: 11 categories	43,668 & 9,367	Y	Oliveira *et al.* [2023]
Play store reviews (2024)	Google Play Store	big six; polarity: positive, negative, and neutral	300	Y	Siqueira *et al.* [2024b,a]
OLID-BR (2024)	Twitter, YouTube, 55chan, news comments	toxicity_type & target_type	6,354	Y	Trajano *et al.* [2024]
HEDOS (2024)	X (Twitter)	toxic, non-toxic, trash	5,492	–	Lima *et al.* [2024]
Multi-label corpus	Twitter (X)	Primary and Secondary emotions defined by Plutchik model	12,160	Y	Mendes *et al.* [2024b]
Translated ISEAR Corpus	ISEAR corpus & augmented version	anger, disgust, fear, guilt, happiness, sadness, shame	7,474 & 13,16 1	Y	footnote 48

[Martinazzo, 2010] on Kaggle as Words in Portuguese (BR) and Emotions dataset[22]. Further this work was extended by extracting synonyms using a dictionary[23] and provided a total of four datasets.

Another corpus based on short news text was built by Dosciatti [2015] for emotion mining (the big six and neutral emotions) and opinion mining (to determine polarity) under the Emoções.BR project and is available on request. A total of 2,000 news headlines with short descriptions, containing on average 23 words, were extracted from global, politics, police, and economics categories. Experts in linguistic and computational linguistics annotated these texts with predominant emotions and their intensity (low, medium, or high). The corpus is highly imbalanced per emotion with distribution: anger (4%), disgust (13%), fear (11%), happiness (9%), neutral (27%), sadness (23%), and surprise (13%). Its balanced version [Dosciatti *et al*., 2013] contains 250 news per emotion but without intensity information. Along with the corpus, a text annotation system, computational tools, and a data balancing approach are described and provided by the author.

Santos *et al*. [2014b] built and analyzed a multilingual emotion mining corpus, which is a less explored area. The authors used reviews of Amazon's Kindle e-books as a source for two reasons: (i) since diverse emotions can be evoked by books, the likelihood of finding different emotions in their reviews is greater; (ii) a wide global customer base offered the possibility of finding reviews in different languages. The corpus contains 150 reviews manually annotated among anger, anticipation, disgust, fear, happiness, sadness, surprise, and trust emotions by two annotators (masters in computing) using an emotion dictionary. A corpus is imbalanced with more instances of happiness emotion. The authors mentioned that (i) agreement between annotators on the presence of emotions is a critical part in the corpus building process; (ii) shorter reviews are less noisy, thus identifying underlying emotion is easier compared to lengthy reviews; (iii) though most of the text analysis resources are English language-specific, they demonstrated that machine translation would be helpful for other languages, and advised to be aware of translation errors. There was no information available on the availability of the corpus.

Santos *et al*. [2014a] performed multilingual emotion mining using parallel corpora. Parallel corpora are created such that the texts in both corpora are translations of each other and structurally identical. For such a corpus, annotating one language can be sufficient. Their parallel corpora consist of original music lyrics in English and their translation in BP obtained from the website[24]. The authors performed emotion mining of a total of 850 stanzas, considering Plutchik's 8 emotions: anger, anticipation, disgust, fear, happiness, sadness, surprise, and trust at the stanza level. Emotion is identified at word level within a stanza and then normalized by the total number of words present in the stanza. The resulting value between 0 and 1 is used to determine the intensity of the emotion as *does not evoke, vaguely evoke, or clearly evoke*. Three annotators annotated it using a dictionary. The authors analyzed this corpus and noted that emotion mining with machine learning outperforms sentiment lexicon; intensity and distribution of emotions depend upon the chosen song, irrespective of the language chosen for the analysis. There was no reference made to the corpus's availability.

Pelle and Moreira [De Pelle and Moreira, 2017; De Pelle, 2019] created the first Brazilian corpus, OffComBR[25], to detect hate speech from online comments. The authors used comments written on politics and sports sections from news websites. They categorized text for racism, sexism, homophobia, xenophobia, religious intolerance, and cursing offenses. Due to annotator limitation, OffComBR-2 contains only 1250 instances where at least two judges were in agreement and its subset OffComBR-3 consists of instances where all judges were in agreement. OffComBR-2 has dominance of cursing offense text. Overall, both versions contain more negative instances. Their main observations include: context is crucial while classifying hate speech; language patterns and its meaning used on the web changes constantly; offensive text was frequent in the cursing category; longer comments were less offensive; and there's a need to discern hate speech without meddling in speech freedom. The authors provided their annotation tool[26] to help other annotators to label their data. Also, they have provided baseline results [De Pelle and Moreira, 2017] and ensemble classifier Hate2Vec to detect offensive comments on the Web [De Pelle, 2019].

Nascimento *et al*. [2018] created the MQDEmotion2018[27] corpus by extracting posts from the social network 'Meu Querido Diário' (MQD). On MQD, users share their daily lives expressing sentiments and emotions; even users can label their posts with the basic emotions. This corpus dataset contains a total of 18,601,010 words and a vocabulary of 265,741 distinct terms. The authors selected a total of 79,523 entries categorizing anger (6,323), disgust (1,512), fear (6,112), happiness (32,672), sadness (27,642), and surprise (5,262) emotions.

Nascimento *et al*. [2019] created a corpus[28] by collecting content from Brazilian imageboard 55chan, and Twitter to detect hate speech with categories offensive or not. Data were filtered using the LIWC 2015 BP lexicon dictionary for anger, anxiety, and positive emotions. Data (3,836 instances) containing words from anger and anxiety categories were selected from the 55chan board and labeled as offensive. Similar amounts of data containing positive emotions were selected from Twitter and labeled as non-offensive. The authors also provided baseline classification measures for the corpus.

Wingenbach *et al*. [2019] developed verbal vignettes in

[22] `https://www.kaggle.com/datasets/antoniomenezes/words-in-portuguese-br-and-emotions`. Accessed on 06 June 2026.

[23] `http://www.academia.org.br/sites/default/files/publicacoes/arquivos/cams-10-dicionario_de_sinonimos_da_lingua_portuguesa-para_internet.pdf`. Accessed on 06 June 2026.

[24] `https://www.letras.mus.br/`. Accessed on 06 June 2026.

[25] `https://github.com/rogersdepelle/OffComBR`. Accessed on 06 June 2026.

[26] `https://github.com/rogersdepelle/hatedetector`. Accessed on 06 June 2026.

[27] `https://github.com/LaCAfe/MQDEmotion2018`. Accessed on 06 June 2026.

[28] `https://github.com/LaCAfe/Dataset-Hatespeech`. Accessed on 06 June 2026.

BP, English, and German languages for emotion elicitation and recognition purposes. Vignettes include the following basic and complex emotions: anger, disgust, fear, happiness, gratitude, guilt, sadness, and neutral. These were written from a first-person perspective, analyzed and rewritten until they reached at least 80% agreement. The vignette script is made available as supplementary material.

Fortuna *et al*. [2019] constructed a hierarchically-labeled hate speech corpus using tweets in Portuguese language (considering all dialects). Their work focused on the hate category among offensive speech with the objective of acquiring discrimination posts based on religion, gender, sexual orientation, ethnicity, and migration. They used Twitter API and R to crawl through the messages with 19 keywords, 10 hashtags, and 29 specific profiles to search for hate speech in the period of January to March, 2017. With the search instances, they built a hate speech lexicon for Portuguese. They filtered duplicated and non-Portuguese language tweets and removed HTML tags and a post less than three words, then anonymized all tweets. Per user maximum 200 tweets were selected for diversity. With these, the final corpus comprises 5,668 posts from 1,156 different users. Non-expert and native Portuguese eighteen students volunteered for binary annotation of posts in hate or no-hate categories with the provided guidelines. Inter-rater agreement with Fleiss Kappa was found as low as 0.17, attributed to difficulties in recognizing the nuances of hate speech by non-experts. 31.5% of collected posts were annotated as a hate speech.

An experienced researcher designed hierarchical class structure with fine-grained 81 hierarchical multiple label schema to comprehend various forms of hate speech and relation among them (the way in which they overlap). Earlier annotated messages as hate were classified into hierarchical multiple labels using open coding methodology. Traversing towards the depth of this structure, the class frequencies vary from larger to smaller (see Table III in [69]). This schema allowed authors to identify less common hate speech categories. Another annotator validated this class structure for 500 messages. Cohen's Kappa agreement within these two annotators was found to be 0.72. This agreement does vary per class, showing difficulty among specific classes to annotate. Subjectivity in annotation is reflected with different inter-annotator agreement per category. The authors provided a baseline binary classification result of 0.78 F1-score for 10-fold cross-validation using pre-trained word embeddings and LSTM models. The authors mentioned the possibility of sampling bias in their corpus due to the nature of the API and the methodology used. Thy made available their corpus[29], labeling schema, and codes[30].

da Silva [2020] built stock market domain corpus[31] based on the Plutchik's Wheel of Emotions considering joy vs sadness, anger vs fear, trust vs disgust and surprise vs anticipation emotions. For this work, the author collected tweets containing at least one of the stock code indexed in IBOVESPA, Brazilian Stock market, for a period of March to May, 2014, resulted in a total of 5,068 non-repeated tweets. These tweets were manually annotated using a crowd-sourcing method. The author created a mobile app and guidelines for annotation process and on stock market slangs. By declaring lottery type prize, the author found 442 volunteers of age 18-60 years of which more than 50% obtained graduation or higher degree. Among them, 14.5% participants had trading experience in stock market and 14.5% participants had technical analysis knowledge. The author considered tweets annotated by at least three participants and assigned emotion by two of them was considered as a label. A total of 4,553 tweets were annotated by at least one participant. The final corpus contains annotated tweets per emotion pair as 4,149 for trust and disgust; 4,138 for joy and sadness; 4,170 for anticipation and surprise; and 4,158 for anger and fear.

With the objective of building a larger corpus, the author defined hashtags related to joy, sadness, trust, disgust, anticipation and surprise emotions. Regardless of domain, tweets in Portuguese language with one of the defined hashtag were collected during September, 2015 to October, 2016. This free domain corpus contains 230857 tweets with following frequency and emotion: sadness (109,768), joy (71,004) , fear (27,488), disgust (9,207), trust (5,192) ,surprise(5,049), anger (4,006), anticipation (1,326) where emotions annotated using distant supervision method. No information was mentioned on its availability.

Leite *et al*. [2020] created a large-scale toxic language corpus for Brazilian Portuguese, ToLD-Br, to detect toxic speech by analyzing Twitter posts. The authors collected posts using GATE Cloud's Twitter Collector, for the period July-August 2019 with two strategies. In the first strategy, tweets were searched for predefined BP hashtags and keywords. In the other strategy, tweets were collected from 50 influential users without any keywords. They found more than 10M unique posts. Among them, 21K posts were selected randomly for annotation, of which 60% posts were collected using keywords. All posts were pseudo-anonymized by replacing the name of the person by the word "user".

For the annotation, 42 volunteers (age 18–37 years) with diverse demographic profiles conforming with the Brazilian Institute of Geography and Statistics were selected by public consultation at the Federal University of Sao Carlos, Brazil. The authors provided demographic information of annotators along with their corpus. Each post is annotated by 3 volunteers, and each annotated 1500 posts among LGBTQ+phobia, obscene, insult, racism, misogyny, and/or xenophobia, otherwise none. Next, all posts were labeled as toxic or non-toxic. If any of the annotators flagged the post on an offensive category, it was labeled toxic. Inter-rater agreement was the most for LGBTQ+phobia, suggesting it has more definitive lexicons. Lower agreement was found for obscene and racism classes. Lexicons used to classify obscene and insult were overlapped showing the confusion within the volunteers for the annotation.

The authors provided a baseline macro-F1 score of 76% using BERT model for monolingual binary classification with 80-10-10% data split into training-development-

[29] `https://github.com/paulafortuna/Portuguese-Hate-Speech-Dataset`. Accessed on 06 June 2026.

[30] `https://github.com/paulafortuna/SemEval_2019_public`. Accessed on 06 June 2026.

[31] `https://www.kaggle.com/datasets/fernandojvdasilva/stock-tweets-ptbr-emotions`. Accessed on 06 June 2026.

test sets. Their multilabel analysis pointed that the model misclassified less frequent toxic categories compared to others, suggesting more data is necessary for the model to learn subtleties. Monolingual model performance outperforms transfer learning and multilingual model, demonstrating how large-scale monolingual corpus is necessary to build a reliable model.

Cortiz *et al*. [2021] created the first TER corpus[32] classifying texts in 28 fine-grained categories in BP by reviewing emotion categories defined in the GoEmotions dataset. GoEmotions[33] corpus [Demszky *et al*., 2020] contains 58,000 carefully curated Reddit comments in English, which are manually annotated in 27 different fine-grained emotions. During the first review stage, seven researchers from different backgrounds suggested BP translation of emotions from English based on the definition of each emotion and created an emotion list. These emotions were reviewed for consistency in the second stage, resulting in 28 distinct emotions [see Table I]. The authors further developed a set of lexical items for these emotions. Using this set, the authors collected and annotated 47,405 tweets with a weak supervision approach. The authors provided baseline results using fine-tuned pre-trained bidirectional encoder representations from transformers (BERT) models and suggested that a weak supervision approach works better for TER in BP, which is a low-resource language. This is also available on Kaggle as Go Emotions PT-BR dataset[34].

Around the same time, Hammes and de Freitas [2021] worked on filtered version of the GoEmotion dataset [Demszky *et al*., 2020] by translating it into Brazilian Portuguese and then performed sentence-level analysis using transfer learning with the BERTimbau model [Souza *et al*., 2020]. The BERTimbau model is based on the transformer architecture and is developed for BP language. It has achieved state-of-the-art performance in many NLP tasks. The authors noted that not all translated sentences produced correct idiomatic (or figurative) expression and thus affected the emotion recognition results. The data and code resources are available[35].

Souza [2022] used the comments in the discussion forums of the massive open online course (MOOC) for TER analysis. The author used a mining tool built in Python, considering the big-six emotions but replaced disgust with dislike, as it is more common among students. The training set contains 538 phrases with the following emotion distribution: anger (84), dislike (90), fear (84), happiness (112), sadness (84), and surprise (84). The test set includes 228 phrases: anger (36), dislike (36), fear (36), happiness (45), sadness (36), and surprise (36). The author believes that knowing students' emotions will be helpful to MOOC teachers for planning their strategies to encourage students to complete the course.

During COVID-19 pandemic when The Brazilian Health Regulatory Agency (Anvisa) authorized the emergency use of vaccines Penteado *et al*. [2022] performed an analysis to identify primary emotion trending among twitter users. Total 1,644,508 corresponding posts were collected using keywords related to vaccine and Anvisa for a period of 24 hours on 17-18 January, 2021. Among these most retweeted 100 posts were selected to identify influential contents and their/respective underlying emotions by adopted method and 85 of them contained some emotions. From the most to least prevalent emotions observed were indignation, enthusiasm, happiness, pride, suspicion, cynicism, defiance, hope, anger, envy, empathy, disgust, and sadness. The authors noted that the posts revealed positive responses over vaccine approval. No information on resource availability was found.

Vargas *et al*. [2022] constructed a large hate speech corpus, HateBR, which is manually annotated by experts for precise classification. This is constructed by collecting comments from public Instagram accounts of six Brazilian politicians (4 female, 2 male) of two different political parties for the period August 2020 – January 2021 using Instagram API. From each of the six accounts, comments were collected from the most popular post. Total 30 posts and 15000 comments were collected. In the preprocessing step, comments with irrelevant text like links, characters without any semantic values and made solely on emoticons were discarded, providing a total of 7000 comments. For the annotation process, the authors defined a new annotation schema that can autodetect offensive and hate speech precisely. To annotate, experts with doctoral degrees (completed or pursuing) in linguistics, hate speech, and computer science with diversity in their political opinion, color, and region were selected to avoid bias.

The annotation was done in 3 ways: (1) binary classification as offensive or non-offensive containing 3500 comments in each category; (2) further offensive comments were classified as per their intensity into high (778), moderate (1044), and weak (1678); (3) Among 3500 offensive comments, 727 comments which provoked violence or hate against any group were classified into 9 hate speech groups: xenophobia, racism, homophobia, sexism, religious intolerance, partyism, dictatorship apologists, antisemitism, and fatphobia. Remaining 2773 comments were classified as non-provoking. Fleiss' Kappa rate was 74% for highly offensive comments and 46% for moderate intensity. The authors provided a baseline result of 85% F1-score with SVM classifier that outperformed earlier state-of-art results for the Portuguese language. No statement regarding corpus availability was given.

Rodrigues *et al*. [2023] built a corpus by collecting tweets related to Brazilian soccer team Palmeiras using Twitter API. They collected 100 tweets per 15 minutes during the matches the team played. Taxonomy was created for support, anger, happiness, and funny emotions and implemented the logic with NLP++ language on HPCC system which is open-source big data analysis platform. The logic is described in the paper. After data analysis, the model was tested with freshly collected tweets during team matches to compare human versus machine emotion recognition. The

[32] `https://github.com/diogocortiz/PortugueseEmotionRecognitionWeakSupervision`. Accessed on 06 June 2026.

[33] `https://github.com/google-research/google-research/tree/master/goemotions`. Accessed on 06 June 2026.

[34] `https://www.kaggle.com/datasets/antoniomenezes/go-emotions-ptbr`. Accessed on 06 June 2026.

[35] `https://github.com/Luzo0/GoEmotions_portuguese`. Accessed on 06 June 2026.

data and codes resources are available[36].

de Oliveira Moraes [2023] collected 5571 tweets about Brazilian touristic city Campos do Jordão using twitter API for a period 2 October to 31 December, 2022. Among these, 2184 posts from 2 October, 2022 to 8 November, 2022 were manually labeled by the author based on predominant emotion. As it is a tourist place, the author created a new emotion class as desire, new two categories of propaganda and information, and due to less instances, anger and fear instances labeled as aversion. The author mentioned the distribution of 4183 posts as None i.e. no emotion (35%), advertising (12.6%), information (16.1%), and 36.3% posts classified among aversion, happiness, sadness, surprise, disgust, and desire. The author noted that labeling was challenging as posts contained mixed emotions, writer' perception as well as towards the city and multi-label strategy would have yielded better understanding instead of choosing dominant emotion. No information is given about resource availability.

Silva *et al*. [2023] built a corpus by collecting tweets posted by official news agencies around the election period in Brazil from June 24, 2018 to April 1, 2019 to assess their support for pre-candidates and candidates running for office. The authors considered Twitter accounts of print and online newspapers, TV and radio channels, magazines, internet portals, etc., which at least 100 followers and posts. They chose a total of 35 accounts to collect daily tweets using Twitter API. A total of 53,932 posts were collected and most of them contained news headlines. The posts were preprocessed to store in MySQL database. The authors identified subjectivity, polarity, and big six emotions per new agency' posts per candidate. They found a prevalence of sadness, surprise, and fear in headlines. Their corpus and source codes are available on request.

Fernandes-de Oliveira *et al*. [2023a,c] analyzed public posts on Facebook and Instagram to find sentiments and emotions of Brazilians regarding COVID-19 vaccine and vaccination process during January 1st, 2020 and December 31st, 2021. To collect data, the terms related to COVID-19 vaccine in BP were searched on public-fan pages with the help of CrowdTangle graphical interface. A total of 2,965,376 and, 974,951 posts were found on Facebook and Instagram respectively. To avoid biases due to high engagement and platform-specific algorithms randomly, 1067 posts were chosen for the study from both platforms. These posts were sorted into three groups: no emotion expression, emotion expressed and identified, and unidentified emotion. Only posts with identified emotions were considered for further processing. Thus, 523 and 620 posts were selected respectively from Facebook and Instagram.

Next, emotion identification and annotation were carried out manually. A total of 56 emotions were defined (Table 1 in [Fernandes-de Oliveira *et al*., 2023a] & [Fernandes-de Oliveira *et al*., 2023c]) and represented on valence-excitation axes. While categorizing the post and annotating emotions, the authors scrutinize the context per post. Among the 56 descriptors described, 35 emotions were detected from Facebook posts and 30 emotions from Instagram posts. Prevalent positive emotions were trust, interest and hope, whereas negative emotions were worry, disapproval, and concern. Negative emotions were found to be contextual to social and political scenarios. There is no reference made to the corpora's availability.

Weitzel *et al*. [2023] created a hate lexicon and an annotated corpus[37] for hate speech at sentence level in BP. The authors collected 80K tweets during the Brazilian election period in 2018–2019 using Twitter's API and Tweepy python library using default hate or offensive keywords. With these keywords, they developed the HALEX (HAte LExicon) lexicon with 512 hate words. Knowing the Brazilian habits of using swear words in happiness or hatred expressions, they removed the swear terms from HALEX. All posts were preprocessed to remove mentions, emojis, URL, special characters, repeated characters, digits, punctuations, and duplicate posts. Then, all abbreviations were normalized and tweets with less than three tokens were discarded. With HALEX, preprocessed 33,771 tweets were classified into hate (5,575) and non-hate categories (28,196) by developing a Python routine. This classification was manually investigated by four native BP researchers (21–60 years). The author mentioned that they found successful inter-rater agreement with Cohen' Kappa metric. As this binary corpus is imbalanced, the authors undersampled the majority class and provided a baseline result by performing experiments with various models.

Oliveira *et al*. [2023] created the hate speech corpus, TuPy[38], and extended it to TuPy-E[39], by including three existing corpora–ToLD-Br [Leite *et al*., 2020], HateBR [Vargas *et al*., 2022], and of Fortuna *et al*. [2019], it is also available on GitHub[40]. The authors collected tweets from May–July 2023 with Twitter API by filtering for offensive keywords defined in earlier studies and also collected randomly without any keywords. The authors chose the same number of instances with and without keywords to stratify them. Preprocessed data to remove duplicate, metadata, and external links. Then, they anonymized the tweets and converted to lowercase. Before incorporating three existing corpora into their data, the authors processed the corpus by Fortuna et al. [69] and ToLD-Br with majority-vote for better interoperability.

For annotation, the authors chose nine candidates with master and doctoral education background, with different gender, ethnicity, and political orientation. The corpus of 43,668 instances was annotated for aggressive (12,547) and non-aggressive (31,121) language. Then, aggressive comments were classified into hate speech (9,367) and not hate speech (3,180) categories. Finally, hate speech instances were labeled among ageism (57), aporophobia (66), body shame (285), capacitism (99), LGBTphobia (805), political (1149), racism (290), religious intolerance (108), misogyny (1,675), xenophobia (357), and other (4476) such that a com-

[36] `https://github.com/pedro-lima-rodrigues/EmotionsDetection/`. Accessed on 06 June 2026.

[37] `https://github.com/LuanPCunha/TCC`. Accessed on 06 June 2026.

[38] `https://huggingface.co/datasets/Silly-Machine/TuPyE-Dataset`. Accessed on 06 June 2026.

[39] `https://huggingface.co/datasets/Silly-Machine/TuPyE-Dataset`. Accessed on 06 June 2026.

[40] `https://github.com/Silly-Machine/TuPyE-Expanded-Brazilian-Hate-Speech-Dataset`. Accessed on 06 June 2026.

ment may have more than one label. Based on statistical analysis, the authors pointed out the subjectivity in the annotation process and the influence of possible biases. They provided base results with BERTimbau large models and also performed various analyses with N-gram and graphical representation. Their codes are available publicly on HuggingFace and GitHub.

Siqueira *et al*. [2024b,a] developed a corpus[41] of polarities and emotions by collecting reviews in BP from the first top ten most downloaded applications on Google Play Store in May 2023. The selected apps were from different categories such as finance, social networking, utilities, and photo editors. The data collection process was carried out using Python and google-play-scraper2 library. A total of 3000 reviews were manually classified and annotated for polarity among positive, negative and neutral, and for emotions among anger, disgust, fear, happiness, sadness, and surprise. All annotations were finalized by three annotators unanimously for accuracy and reliability. The authors noted that the proportion of negative emotions is much more than positive ones as 952 comments were classified as disgust and surprise class contained only 4 comments.

Trajano *et al*. [2024] built the span level labeled corpus named Offensive Language Identification Dataset for Brazilian Portuguese (OLID-BR) to detect offensive language. OLID-BR is prepared to perform five different NLP tasks as 1) is_offensive: is the comment toxic or not; 2) offensive_type: if toxic then label it among following categories: health, ideology, insult, LGBTQphobia, other-lifestyle, physical aspects, profanity/obscene, racism, religious intolerance, sexism, and xenophobia; 3) is_targeted: is the comment targeted; 4) targeted_type: if yes, then choose the type of target among individual, group, or other; and 5) toxic_word_list: provide toxic span, i.e, list toxic words from the comment.

To encompass the diverse user behaviors and to avoid data-source bias, the authors collected data from two different social-media platforms, Twitter and YouTube. From Twitter, tweets were collected using two methods. In the first, posts were selected from public figures from politics, news, enterprise, artists, influencers, sports, and entertainment categories. In another, tweets were collected by filtering for toxic keywords using Twitter API. The same strategy is used to select videos from YouTube, where comments and replies to those comments were collected. The authors manually selected videos with the possibility of acquiring offensive comments which have a high number of views, comments, and controversial or provoking contents in the title or description. The authors also used raw text as data source from earlier published corpora in Portuguese language which have different annotation schema than theirs: OfComBR [Nascimento *et al*., 2018], ToLD-Br [Leite *et al*., 2020], and corpora by Nascimento *et al*. [2019] and Fortuna *et al*. [2019].

The authors identified toxic posts using Perspective API with threshold $\geq$0.7. Then, comments were processed to remove duplicates, non-BP words, misunderstood texts, and to anonymized usernames, hashtags, and URLs. They created a detailed guideline[42] to train annotators who were selected with predefined criteria. The annotators performed five NLP tasks stated above. To finalize the annotation, the authors adopted a majority vote to label as is_offensive, is_target, and targeted_type. For the multilabel offensive category, the less restrictive criterion of 'at least one' was adopted. The annotation process was carried out in iterations to refine the annotation process. With each iteration, inter-rater agreement in terms of Krippendorf's Alpha of toxicity labels was improved, and it achieved moderate agreement. The authors reasoned that as the task is subjective and annotators varied age, sex, and educational background might have led to high disagreement.

The corpus (total 6,354 instances) is available in CSV and JSON format with a training set containing 4765 instances and a public test set containing 1,589 instances for task A-D. It contains toxic span and metadata for each comment. Also, the authors provided an addition_data.json file that contains 7,184 comments with incomplete annotations that require careful investigation to confirm its reliability. Another private test set of 1,589 instances with complete annotation is guarded for future use like competition. The corpus is moderately to highly imbalanced per label. The authors provided baseline precision, recall, and F1-score for classification tasks using pretrained BERT model and performed named entity recognition (NER) task for toxic span detection. The schema of OLID-BR is compatible to resources in many other languages, making it potential to use for multilingual/cross-lingual tasks. Thus, this corpus is useful for supervised, semi-supervised, and automated content moderation to filter offensive language while presenting contextual information through the previously mentioned toxicity type-target-span detection. It is available on Kaggle[43] and HuggingFace[44]. Also, its source code[45] is available.

Lima *et al*. [2024] created a hate speech corpus with xenophobia as a main theme titled hate discourse in online spaces (HEDOS). The authors collected posts from X (formerly Twitter) for the period before, during, and after the 2022 election in Brazil using *snscrape* library and TwitterSearchScraper method. Username, mentions, and origin of the data such identification references were removed. They created a dictionary of Brazilian gentilic, obscene, or inappropriate words in their singular and plural form to preprocess those posts. For the annotation process, they defined the hate speech categories with three labels, toxic, non-toxic, and trash, with their respective proper meaning to help the annotators.

With context, these definitions were kept flexible for proper annotations. Also, they provided the context to the posts when swear words were used and for comments related to events other than elections. Three computer science students (2 males, 1 female) from Federal University of Pelotas (UFPel) with different social, demographic, and

[41] `https://zenodo.org/records/10823148`. Accessed on 06 June 2026.

[42] `https://dougtrajano.github.io/olid-br/annotation/guidelines.html`. Accessed on 06 June 2026.

[43] `https://www.kaggle.com/dougtrajano/olidbr`. Accessed on 06 June 2026.

[44] `https://huggingface.co/datasets/dougtrajano/olid-br`. Accessed on 06 June 2026.

[45] `https://dougtrajano.github.io/olid-br/`. Accessed on 06 June 2026.

regional backgrounds annotated the corpus. The annotation with at least two inter-annotator agreements was accepted. HEDOS contains a total of 5,492 posts with 4,301 toxic, 1,190 non-toxic, and 1 post as trash. Among toxic posts, most of them are related to regionality, indicating xenophobia. Thus, HEDOS represents xenophobia as a main category among hate speech. The Fleiss's kappa metric was just 21%, indicating poor inter-annotator agreement, thus having low reliability in terms of data integrity. No information was given about HEDOS availability.

Mendes *et al*. [2024b] built a multi-label corpus[46] considering primary and secondary emotions defined by the Plutchik model. The authors defined synonyms for the desired emotions and collected 12,160 posts from social network X using Python web scrapper. After data analysis with SVM and BERT, they expressed a need to develop a more linguistically diverse resource by considering a variety of sources.

With an effort of many large groups of psychologists all over the world and student respondents, the international survey on emotion antecedents and reactions (ISEAR) dataset[47] was developed under the ISEAR project [84]. In a survey, respondents addressed the situation they appraised and reacted with the following emotions: anger, disgust, fear, guilt, happiness, sadness, and shame. Machine translation of the ISEAR dataset in BP is made available as ISEAR Corpus Translated to Portuguese BR[48] by the Kaggle user. Additionally, the translated dataset was augmented using a synonym resource.

4.4 Multimodal ER Data Resources

CH-Unicamp expressive viseme images database [Costa, 2015] is an audio-visual-motion capture corpus constructed to develop 2D facial animation synthesis methodology incorporating head and face movements to mimic the various emotional states. This corpus was recorded with 10 BP native professional actors (4 females, 6 males) under controlled conditions with predefined text containing all BP phonemes and context-dependent visual phonemes with 22 emotions defined by the OCC model and neutral. Another experiment was designed with the same predefined utterances to study how different personality traits modulate face emotion expression but with the big six, shy, neutral, and extrovert expressions. For the motion capture session, 63 reflective markers were spread out over the face and head of the participants. This corpus comprises 782 RGB uncompressed facial images of size 1920×1080 pixels representing 34 expressive visemes. With this, a CSV file details all metadata of every image including emotion, phonetic details with context, context-dependent viseme, and coordinates of feature points. The complete set of text excerpts used to record each emotional state, the expressive speech face model, analysis, and code are provided. A perceptual evaluation of valence, emotion, and visual cues was performed by 48 native Brazilians (19–66 years), who were students and staff from University of Campinas. This resource is available on request.

Gonçalves *et al*. [2017] conducted an experiment to assess participants' emotion while playing a video game by measuring their behavior tendencies, speech, and facial expressions using multiple sensors. The author obtained the approval from the ethical committee to carry out the experiment. Thirty users (80% male) in the age group 18–32 years with varied levels of education and video game playing experiences participated. Camera and microphone recorded facial and vocal expressions during the game interaction, which lasted for 10 minutes. Logical sensors were used to assess behavioral tendencies, where 12 game actions were mapped to measure the performance. Finally, the users informed their subjective emotion through self-assessment manikin (SAM). The authors devised a procedure to represent acquired emotions from different sensors and SAM to octants on Scherer's emotional semantic space. They observed differences in the mapped emotions with SAM. Three psychologists independently mapped the sensors' data on octants and provided combined representation per user. The authors compared the octants obtained from sensors and those created by psychologists using ANN. No information is provided on data availability.

The Brazilian audiovisual dataset of emotion expressions (BADEM, [Carlos, 2022]) is a part of the large project *Pesquisa de processo resultado em psicoterapia: um estudo computacional* with an objective to advance BP-ER modeling. It is recorded with 12 native Brazilian actors from different states (6 females and 6 males) using 12 predefined phrases of neutral content. The author obtained the approval from the ethical committee to carry out the experiment. The BADEM contains a total of 1,008 audio-videos with the big six and neutral emotions. The duration of each recorded phrase per emotion is approximately 4 seconds. Eight clinical psychologists with professional experience in a range of 3–20 years evaluated the naturalness of performing emotions on a 5-point Likert scale, from *not at all natural* to *completely natural*. To construct the database, the authors adopted the process of well-known foreign language datasets. With agreement over 80% among evaluators and around 90% of videos were rated from moderately natural to completely natural, authors noted that their objective of creating a possibly natural ER dataset is achieved with some limitations.

Fernandes-de-Oliveira et al. analyzed videos published on TikTok [Fernandes-de Oliveira *et al*., 2023b] and Kwai [Fernandes-de Oliveira *et al*., 2024] platforms, during the years 2020–2022, to learn the emotions of Brazilians towards COVID-19 vaccine and vaccination process. Videos with the hashtag "vacina" along with metadata were extracted using Python TikTokAPI library, considering interactivity ranking provided by the platform. The resulting videos were manually processed for BP language and COVID-19 vaccine related theme. The processed videos were further investigated for emotion expression and identification, which provided 435 videos to compose a final dataset. From the Kwai platform, videos were collected with COVID-19 and vaccine related hashtags using Python from the video watch page of Kwai' graphical interface. Considering BP language, public availability, removal of duplicates, and where emo-

[46] `https://github.com/MiningEmotion/EmotionsMiningPTBR/`. Accessed on 06 June 2026.

[47] `https://www.unige.ch/cisa/index.php/download_file/view/395/296/`. Accessed on 06 June 2026.

[48] `https://www.kaggle.com/datasets/antoniomenezes/isear-corpus-translated-to-portuguese-br`. Accessed on 06 June 2026.

Table 4. Multimodal Emotion Recognition Resources in Brazilian Portuguese
Mode: A (audio), Img (image), M (motion), S (Sensors), V (video), T (text); Status: Y (yes), R (upon request), N (No), – (no information)

Resources (year)	Source	Emotions/Affect	Mode (Number)	Status	Ref
CH-Unicamp (2015)	Recorded	22 emotions from the OCC model and neutral Personality traits: big six, shy, neutral and extrovert	Img (782), T	R	Costa [2015]
Game interaction (2017)	Video Game	octants on Scherer's emotional semantic space	A, I, S	–	Gonçalves *et al.* [2017]
BADEM (2022)	Recorded	big six and neutral; naturalness categorized on 5 point Likert scale	A, V (1008)	R	Carlos [2022]
COVID-19 vaccination (2023)	TikTok	34 emotions	A, V, T (435)	–	Fernandes-de Oliveira *et al.* [2023b]
COVID-19 vaccination (2024)	Kwai	26 emotions	A, V, T (355)	–	Fernandes-de Oliveira *et al.* [2024]

tion is expressed and identified, a total of 355 videos were detected. Taking into account audiovisual characteristics, text transcripts, and context of videos from both platforms, emotions were annotated using predefined descriptors (Table I in Fernandes-de Oliveira *et al.* [2023b] & Fernandes-de Oliveira *et al.* [2024]).

Also, the TikTok videos were sorted into personal, informative, infotainment, humorous, marketing, and unidentified classes based on modality and content. The Kwai videos were sorted similarly in the first five classes. Among the 56 described descriptors, 34 and 26 emotions were detected from multimodal analyses of TikTok and Kwai videos respectively. As the authors obtained data from Facebook [Fernandes-de Oliveira *et al.*, 2023a] and Instagram [Fernandes-de Oliveira *et al.*, 2023c] posts, for TikTok and Kwai platforms, negative emotions of doubt and disapproval were prevalent and were mostly from humorous, informative, or infotainment category posts. Also, these negative posts were published in social or political context. Similarly, positive emotions of contentment, enthusiasm, and trust were prevalent in the context of vaccines. No information is provided on data availability.

5 Discussion

A data resource is a key element in building ER models that actuate affective interactive systems. The accuracy of the trained model is governed by the language-cultural differences present in the data resource. Therefore, it is widely realized that the creation of a data resource needs to progress both in terms of volume (quantity) and quality across different modalities and encompassing linguistic and cultural diversity.

In this exploratory review, we included studies where defined keywords appeared in the title or abstract and are useful for computation purposes. For example, many works focusing on emotion perception and behavioral analysis have not been included. We explored published work up to 2024 that has open access. The review started with an intriguing curiosity about how many ER data resources are available in speech, facial expressions, and text modalities and assessed

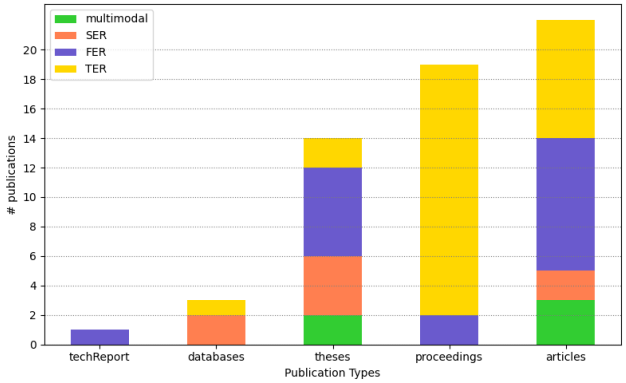


Figure 1. Type of publications per modality.

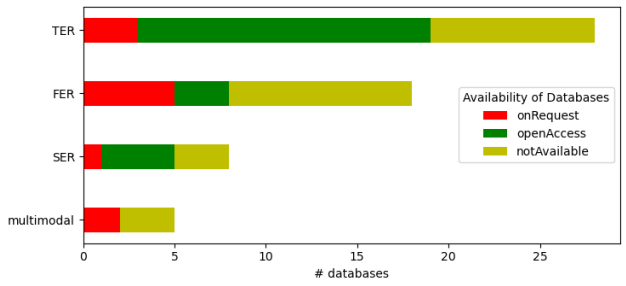


Figure 2. Availability of data resources per modality.

their current status in BP. Among the searched literature, Figure 1 shows the publication types where we found the data resources – 22 articles, 20 conference papers or proceedings, 14 theses or dissertations, 1 technical report, and three data resource repositories across the modalities considered.

We found eight, eighteen, and twenty-eight data resources respectively in speech, facial expression, and text modalities. In addition, we found five multimodal data resources. Of these, less than 60% are either readily available or on request. Figure 2 presents this information graphically. These data resources are listed in Tables 1–4 in the respective subsections 4.1 to 4.4 per modality in section 4.

Emotions explored in all these data resources are listed in tables, and visual representation of their frequency is pre-

Figure 3. Emotions explored in all data resources.

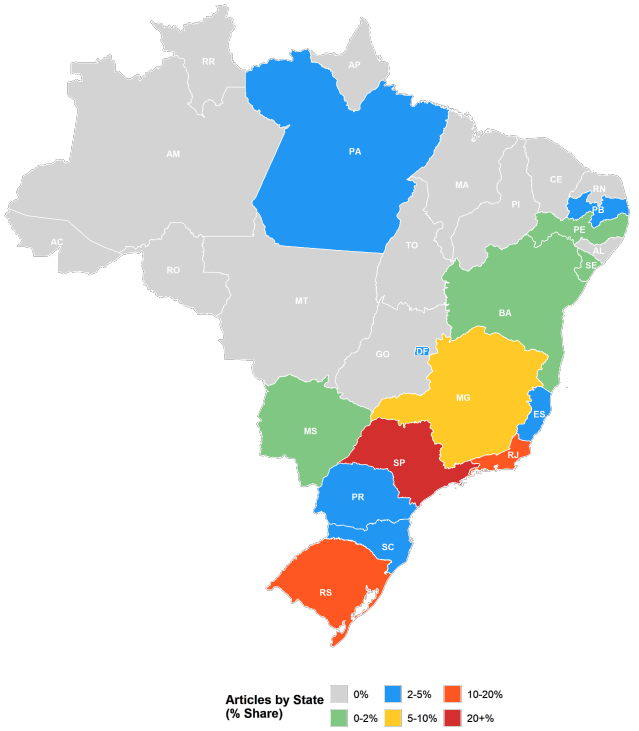
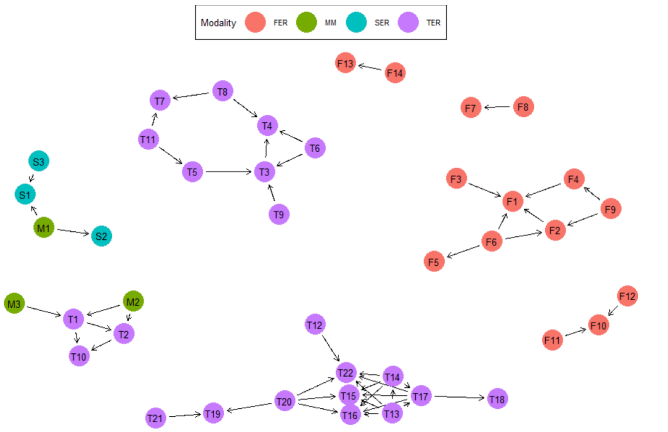


Figure 4. Geographical distribution showing the state wise contribution (in %) of research groups building BP-ER data resources.

sented in Figure 3. Though these data resources are handful and smaller in size, it is interesting to see how their construction and sources have evolved. An intriguing observation is that most of these resources followed categorical emotion model. And in that, mostly concentrated among six basic emotions. The reason can be attributed to the complexities involved in defining, mapping, obtaining, and validating emotions [Fathalla, 2020].

To gain insight into collaboration among different research groups, we intended to look at the geographical distribution of the studies by mapping the state wise participation of research groups within Brazil. The map presented in Figure 4 shows the contribution from the state of São Paulo (SP) to be greater than 20% , followed by 10-20% from the states Rio Grande de Sul (RS) and Rio de Janeiro (RJ), while the states Minas Gerais (MG), Bahia (BA), Pará (PR), Paraná (PR), Santa Catarina (SC), Pernambuco (PE), Sergipe (SE), Paraíba (PB) contribute less than 10%. Overall, the distribution is skewed. International participating institutions are from Canada, USA, Netherland, Portugal, UK, and Spain.

Observing a large variation in the number of available



S1=Torres et al., 2018; S2=Rosa, 2017; S3=Marcacini et al., 2022;, F1=Romani-Sponchiado et al., 2015; F2=Donadon et al., 2019; F3=Negrao 2019; F4=Negrao et al. 2021; F5=Novello et al., 2018; F6=Fabricio 2023; F7=Goulart et al., 2019a; F8=Goulart et al., 2019b; F9=Morais et al., 2024; F10=Vieira 2017; F11=Siqueira et al., 2018; F12=Siqueira et al., 2023; F13=Guimaraes 2014; F14=Nascimento 2016;, T1=Fernandas de Oliveira et al., 2023a; T2=Fernandas de Oliveira et al., 2023c; T3=Dosciatti et al., 2013; T4=Martinazzo 2010; T5=Hammes and Freitas, 2021; T6=Silva 2020; T7=Santos et al., 2014b; T8=Santos et al., 2014a; T9=Rodriguese et al., 2023; T10=Penteado et al., 2022; T11=Mendes et al., 2024; T12=Nascimento et al., 2018; T13=Lima et al., 2024; T14=Oliveira et al., 2023; T15=Fortuna et al., 2019; T16=Leite et al., 2020; T17=Vargas et al., 2022; T18=Guimaraes 2020; T19=Nascimento et al., 2019; T20=Trajano et al., 2024; T21=Weitzel et al., 2023; T22=Pelle and Moreira, 2017;, M1=Carlos 2022; M2=Fernandas de Oliveira et al., 2024; M3=Fernandas de Oliveira et al., 2023b.

Figure 5. Network of citations in modalities shown using directed graph. Nodes represent the studies in the respective modalities and the direction of an arrow shows the studies being cited.

data resources among the modalities (Figure 2), we intended to view the depth of the research landscape in terms of interconnectedness within and across modalities. Figure 5 shows a directed graph in the form of visualizing a network of citations among the available data resources. Up to the time window for our search in our survey, i.e, 2024, the SER, FER and TER modalities are mutually exclusive. Within the SER modality we found only one citation among eight works. Whereas in FER and TER, researchers were aware of some earlier works. In FER modality seven out of eighteen and in TER thirteen among twenty-eight were cited. Within multimodal works none cites the other but refers to SER and TER works.

Table 1 lists the details of SER data resources discovered in our survey. The list spans from a very small corpus of speech acts to the latest one built for ER purposes using speech recognition corpus. Similarly, the objective underlying their construction varies from associating emotions with linguistic and paralinguistic characteristics to finding vocal and speech parameters to distinguish emotions. Sources utilized in resource building present creativity and diversity from lab recording, online contents, operator communication, TV-Radio shows to using BP speech/text recognition corpora for ER purpose, which is a complex task. The number of emotions categorized per data resource varies from a minimum of one, binary (neutral, non-neutral) to seven (big six and neutral). Among the available five (one on request) resources, the largest is VERBO, which is an acted corpus with 7 distinct emotions and a total of 1,167 audios.

The next is, spontaneous corpus CORAA, with only neutral and non-neutral emotions. The organizers of PROPOR 2022 made available CORAA for the contest. They [Marcacini *et al.*, 2022] described the motivation behind its creation and presented an overview of submitted work, but seems the authors were unaware of existing earlier spontaneous data resources. Although VERBO is referred to but is not mentioned explicitly as one of the SER resource. An-

other observation is the researchers participated in the contest mentioned the lack of resources in BP and none were aware of earlier resources. Marcacini *et al*. [2022] explained the motivation behind creation of CORAA as the need of spontaneous BP-SER resources for robust ER models but did not explain how only non-neutral class with male and female voice distinction can improve SER models.

Though not listed in the tables, we included studies Zampar [2010]; Oliveira *et al*. [2014] that used Radio programs for emotion study and Silva *et al*. [2020] that used a private customer care dataset with no primary objective of compiling a data resource. Surprisingly, hardly any of these resources have been cited in the other resource building work, although they discuss the scarcity of resources, except Marcacini *et al*. [2022] mentioned VERBO [Torres *et al*., 2018] and one multimodal work [Carlos, 2022] cited the VERBO [Torres *et al*., 2018] and dataset by [Rosa, 2017] (see Figure 5). This fact of invisibility of earlier resources underlines the importance of this review.

Table 2 lists the FER data resources. The earliest dataset FEI contains only neutral and smiling faces but provides a wider profile view from frontal to $\pm 180^{\circ}$, while the rest provides only frontal view. ER from non-frontal faces is challenging and currently a hot topic in research. FER resources covers age groups from infants to youth to adults, still images to videos, and among them four are spontaneous. These resources have explored emotions beyond the big six, including discrete and dimensional models, complex emotions and emotions at different intensity levels. In addition to audio-visual stimuli, emoticons, video games, and social robots were used for inducing emotions. Facial thermal images, facial landmarks and their mapping to facial action units have been explored for building resources. Tejada *et al*. [2022] have created a pan-trans-cultural dataset and have stressed the need for building language-regional-specific resources to improve ER models. The authors have attempted to retain diversity among the participants following the distribution of Brazilian population [Novello *et al*., 2018; Donadon *et al*., 2019; Fabrício, 2023; Morais *et al*., 2024]. Among the available eight (five on request) data resources, the larger one is FEI with 2800 images containing only neutral and smiling faces.

Text modality provides major (28) contributions to ER data resources and is heavily derived from online contents. In terms of availability, TER corpora dominates with sixteen openly available and three available on request as presented in Figure 2 and Table 3. Sources include posts and comments from news, e-commerce applications, websites, social networking and open learning platforms. For eliciting emotions, vignettes were developed. Both basic and complex emotions are explored. More than one third of corpora explored offensive and hate speech. We included a corpus Fortuna *et al*. [2019] which encompasses all dialects of Portuguese, and have been included while constructing Tupy-E [Oliveira *et al*., 2023] and OLID-BR [Trajano *et al*., 2024] corpora. A unique observation in TER corpora is the valuable contributions coming from non-academic sources (e.g., Kaggle, see footnotes 22, 34, 48). The analysis of the machine translated data resource [Hammes and de Freitas, 2021] underlined the importance of veryfying consistency of languge nuances in translation. Table 4 reports five multimodal corpora, of which two nonspontaneous corpora are available on request. Appraisal dimensionalities have been explored in addition to earlier considered emotion models.

Following guidelines pertaining to the ethical and privacy standards, approval from ethical committee, obtaining consent from participants or legal guardians in case of minors, and declaration of researchers responsibility for the use are critical considerations prior to building the resource and its use. The selected studies have followed the guidelines. For resources built from online experiments, to preserve the identity userId or name was anonymized. On the other hand, when conducting experiment in-person mode, complete information related to the experiments, data collection, usage, and distribution terms were informed to all participants and their consent was obtained. When after the experiment, if the participant was unwilling to publish their images, the researchers deleted all corresponding data. Unbiased contents are preferred for elicitation methods. Where participants were minors, due diligence was given in all aspects of experiments and experiments were carried out under supervision of child psychologists and doctors for infants or children.

With this, we take readers' attention to suggestions made in various reviews regarding data resources. In real life, people display complex emotions. Thus, data resources should be large and diverse enough to learn underlying subtleties. Uniform emotion annotation schema can enhance compatibility within data resources. Since emotion expression and perception are relative, multi-label ER would be useful too [Deng and Ren, 2021]. Multimodal data resources are vital to build effective, intelligent multimodal systems. Along with other modalities, text modality has been demonstrated in improving performance [Poria *et al*., 2017; Deng and Ren, 2021]. It is crucial, however, to use all available modalities with spontaneous or nonspontaneous environments and annotate them in discrete and dimensional emotion classes to endow emotional intelligence to machines [Schuller, 2018; Wang *et al*., 2023, 2022]. If a data resource contains acoustic-phonetic information, it will be helpful to various study branches like health sciences (phono-audiology, psychology, psychiatry), linguistics, etc.

To sum up, the number and volume of BP ER data resources across the speech, facial expression, and text modalities are limited. Further, the number of available data resources either having open access or by request is much less. The sources explored for data resource creation present novelty. Overall, the data resources created are inclusive in accordance with Brazilian society in terms of age, gender, education, region, and ethnicity. Apart from ER studies, some of the data resources are useful in neurological, psychiatric, psychological, and physiological studies. The researchers have discussed difficulties encountered when creating, annotating, and validating the data resource. Perceiving emotion without context is certainly challenging [Torres *et al*., 2018], but even with context, professional judges have mentioned difficulties [Santos *et al*., 2014b; Oliveira *et al*., 2023; Trajano *et al*., 2024]. As a result, it is evident that this is an intricate multidisciplinary endeavor. All studies developing these data resources emphasized scarcity, thereby highlighting the necessity of a larger multimodal emotion data resources in

BP.

6 Conclusion

In a first, this review assesses the current status of the available BP-ER data resources for advancing ER models in BP. Our search explores resources from speech, facial expressions, text, and multimodal modalities to compile a list of existing resources, methodology, validation criteria and results, their accessibility, and emotions studied. Analysis of the resulting studies to gain insights into the Brazilian research landscape and its depth revealed a skewed distribution and weak networking among the research groups. A principal observation in our search reveals the scarcity of BP-ER resources despite BP being globally the eighth most spoken language. A low-resource language constrains the development of ER models. This scenario highlights the necessity of not only larger but diversified in-the-wild multimodal emotion data resources in BP.

After examining wide range of literature, we outline the best practices to consider for building resources to advance ER modeling in BP - consider multimodal modality with diverse emotions, leverage the rich Brazilian diversity to include linguistic and cultural nuances; establish clear guidelines on taxonomy and annotation procedures, even to include regional slangs as influence of social media is becoming integral part of the society, generalizability for inter-corpus or cross-corpus analysis, adherence to ethical standards and reporting, increased collaboration among research groups and expansion of the research landscape.

In the future, it will be interesting to explore data resources among various languages, considering their distribution per types and sizes; cultural and linguistic diversity, and their representation in resources; challenges in acquiring data resources; and comparing those with BP-ER resources. Extending it in the context of how these data resources will be useful for interactive systems and improvements achieved by integrating AI agents.

We hope that this review will be resourceful for future BP-ER studies as well as for any low-resource language studies for advancement towards AEI and for progress in affective interactive systems.

Declarations

Acknowledgements

We thank the anonymous reviewers for their constructive comments that improved the clarity and quality of this work.

Funding

This work was partially sponsored by PCI/CTI/MCTI/CNPq Program process number 444303/2018-9 and by the FAPESP Project 2020/07074-3.

Authors' Contributions

JJGR contributed to the conception of this study. NJ is the main contributor and writer of this manuscript. MRB, JP, and JJGR have added further information to the template. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

All data generated or analyzed during this study are included in this review. Table contents are available as complementary materials.

Further relevant information

This review is performed with openly available data resources and their published documentation/research. All sources are cited properly for transparency. We paid attention not to misinterpret characteristics/features or the intended use of data resources. No data collection, article search, review, or manuscript write-up was performed by AI tools.

References

Balage Filho, P., Pardo, T. A. S., and Aluísio, S. (2013). An evaluation of the brazilian portuguese liwc dictionary for sentiment analysis. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*.

Bastos, G. R. G., Tcheou, M. P., da Rocha, H. F., and Pinto, G. J. S. (2021). emouerj: an emotional speech database in portuguese. *1.0.0 [Data set], Zenodo*. DOI: https://doi.org/10.5281/zenodo.5427549.

Bilewicz, M., Soral, W., Świderska, A., Küster, D., Winiewski, M., and Wypych, M. (2025). The emotional drivers of hate speech: Unpacking the central role of contempt in derogatory communication. *Atlantic Journal of Communication*, 33(5):766–784. DOI: https://doi.org/10.1080/15456870.2025.2525790.

Bruckschen, M., Muniz, F., de Souza, J., Fuchs, J. T., Infante, K., Muniz, M., Gonçalves, P., Vieira, R., and Aluísio, S. Anotação linguística em xml do corpus pln-br. Technical report, Universidade de São Paulo, São Paulo, Brazil.

Bucior, L. and Plentz, P. D. M. (2025). Ai, emotions, and interaction: A social matching system with chatgpt-based conversational agents. In *Escola Regional de Aprendizado de Máquina e Inteligência Artificial da Região Sul (ERAMIA-RS)*, pages 25–28. SBC. DOI: https://doi.org/10.5753/eramiars.2025.16382.

Cambria, E., Livingstone, A., and Hussain, A. (2012). The hourglass of emotions. In Muller, V., editor, *Cognitive Behavioral Systems, Lecture Notes in Computer Science, A. Esposito, A. Vinciarelli, R. Hoffmann*, pages 144–157. Springer, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-34584-5_11.

Carlos, L. M. F. (2022). A criação de um dataset audiovisual brasileiro de expressões emocionais. paradigma centro de ciências e tecnologia do comportamento. Master's thesis, Paradigm Center for Behavioral Science and Technology, São Paulo.

Carvalho, F., Junior, F. P., Ogasawara, E., Ferrari, L., and Guedes, G. (2024). Evaluation of the brazilian portuguese version of linguistic inquiry and word count 2015 (bp-liwc2015). *Language Resources and Evaluation*, 58(1):203–222. DOI: https://doi.org/10.1007/s10579-023-09647-2.

Carvalho, F., Rodrigues, R. G., Santos, G., Cruz, P., Ferrari, L., and Guedes, G. P. (2019). Evaluating the 2015 brazilian portuguese liwc lexicon with sentiment analysis in social networks. *CSBC 2019-8th BraSNAM*. DOI: https://doi.org/10.5753/brasnam.2019.6545.

Carvalho, F., Santos, G., and Guedes, G. P. (2018). Affectpt-br: an affective lexicon based on liwc 2015. In *2018 37th International Conference of the Chilean Computer Science Society (SCCC)*, pages 1–5. DOI: https://doi.org/10.1109/SCCC.2018.8705251.

Chavinda, K., Kalansooriya, P., Weerasuriya, T., de Silva, N., Thayasivam, U., Srikandabala, K., Prabagar, K., and Alahakoon, D. (2026). An emotion aware context adaptive machine learning approach for detecting hate speech in social media. *Neural Computing and Applications*, 38(8):268. DOI: https://doi.org/10.1007/s00521-026-12029-8.

Colamarco, M. and Moraes, J. A. D. (2008). Emotion expression in speech acts in brazilian portuguese: production and perception. *Speech Prosody*, 4:717–719. DOI: https://doi.org/10.21437/SpeechProsody.2008-159.

Cortiz, D. and Boggio, P. (2022). Do you know" saudade"? the importance of a cultural and language-based emotion approach for hci. *arXiv preprint arXiv:2204.06119*.

Cortiz, D., Silva, J. O., Calegari, N., Freitas, A. L., Soares, A. A., Botelho, C., Rêgo, G. G., Sampaio, W., and Boggio, P. S. (2021). A weakly supervised dataset of fine-grained emotions in portuguese. In *Anais do XIII Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana*, pages 73–81. DOI: https://doi.org/10.5753/stil.2021.17786.

Costa, P. D. P. (2015). *Two-dimensional expressive speech animation*. PhD thesis, Department of Electric and Computation Engineering, State University of Campinas, São Paulo.

da Silva, F. J. V. (2020). *Cross-domain emotion detection in tweets*. PhD thesis, Instituto de Computação, Universidade Estadual de Campinas.

Dalvi, C., Rathod, M., Patil, S., Gite, S., and Kotecha, K. (2021). A survey of ai-based facial emotion recognition: Features, ml & dl techniques, age-wise datasets and future directions. *IEEE Access*, 9:165806–165840. DOI: https://doi.org/10.1109/ACCESS.2021.3131733.

De Angeli, A., Johnson, G. I., *et al*. (2004). Emotional intelligence in interactive systems. In *Design and emotion*, pages 262–266. Taylor & Francis.

de Oliveira Moraes, R. (2023). Avaliação de emoções no twitter com métodos de ciências de dados. In *XLIII Encontro Nacional de Engenharia de Produção*, pages 1–12. DOI: https://doi.org/10.14488/ENEGEP2023_TN_ST_404_1987_45647.

De Pelle, R. P. (2019). Identificaçaõ de comentários offensivos na web. Master's thesis, in Postgraduation program in Computation, Information institute, Federal University of Rio Grande do Sul.

De Pelle, R. P. and Moreira, V. P. (2017). Offensive comments in the brazilian web: a dataset and baseline results. In *Brazilian Workshop on Social Network Analysis and Mining (BRASNAM)*, pages 510–519. SBC. DOI: https://doi.org/10.5753/brasnam.2017.3260.

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., and Ravi, S. (2020). Goemotions: A dataset of fine-grained emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4040–4054. DOI: https://doi.org/10.18653/v1/2020.acl-main.372.

Deng, J. and Ren, F. (2021). A survey of textual emotion recognition and its challenges. *IEEE Transactions on Affective Computing*, 14(1):49–67. DOI: https://doi.org/10.1109/TAFFC.2021.3053275.

Dias, C. M. M. (2020). *Elementos visuais na propaganda política*. PhD thesis, Neurosciences, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte.

Domingues Aparecido, T. D., Carrillo, A., Camargo, C. Q., and Stella, M. (2025). Benchmarking psychological lexicons and large language models for emotion detection in brazilian portuguese. *AI*, 6(10):249. DOI: https://doi.org/10.3390/ai6100249.

Donadon, M. F., Martin-Santos, R., and Osório, F. L. (2019). Baby faces: development and psychometric study of a stimuli set based on babies' emotions. *Journal of Neuroscience Methods*, 311:178–185. DOI: https://doi.org/10.1016/j.jneumeth.2018.10.021.

Dosciatti, M. M. (2015). *Um método para a identificação de emoções básicas em textos em português do Brasil usando máquinas de vetores de suporte em solução multiclasse*. PhD thesis, Informatics, Pontifical Catholic University of Paraná (PUCPR), Curitiba.

Dosciatti, M. M., Ferreira, L. P. C., and Paraiso, E. C. (2013). *Identificando emoções em textos em português do brasil usando máquina de vetores de suporte em solução multiclasse*. ENIAC-Encontro Nacional de Inteligência Artificial e Computacional, Fortaleza, Brasil.

Ekman, P. and Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124. DOI: https://doi.org/10.1037/h0030377.

Elfenbein, H. A. and Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128(2):203–235. DOI: https://doi.org/10.1037/0033-2909.128.2.203.

Elkobaisi, M. R., Machot, F. A., and Mayr, H. C. (2022). Human emotion: A survey focusing on languages, ontologies, datasets, and systems. *SN Computer Science*, 3(4):282. DOI: https://doi.org/10.1007/s42979-022-01116-x.

Fabrício, D. d. M. (2023). *Construção e Evidências de Validade de um Banco de Faces para Reconhecimento de Expressões Facias de Emoções Básicas na População Brasileira*. PhD thesis, Post-Graduation program in Psychology, Center of Education and Human Sciences, Federal University of São Carlos, São Paulo, Brazil.

Fabrício, D. M., Ferreira, B. L. C., Maximiano-Barreto, M. A., Muniz, M., and Chagas, M. H. N. (2022). Construction of face databases for tasks to recognize facial expressions of basic emotions: a systematic review. *Dement Neuropsychol*, 16(4):388–410. DOI: https://doi.org/10.1590/1980-5764-DN-2022-0039.

Fathalla, R. (2020). Emotional models: Types and applications. *International Journal of Synthetic Emotions (IJSE)*, 11(2):1–18. DOI: https://doi.org/10.4018/IJSE.2020070101.

Fernandes-de Oliveira, G., Massarani, L., dos Santos-

Jr, M. A., Scalfi, G., and Oliveira, T. (2024). The covid-19 vaccine on the short-video platform kwai: A study of the emotions expressed in brazilian content. *Cultures of Science*, 7(3):166–183. DOI: https://doi.org/10.1177/20966083241280684.

Fernandes-de Oliveira, G., Massarani, L., Oliveira, T., Scalfi, G., and dos Santos-Jr, M. A. (2023a). The covid-19 vaccine on facebook: A study of emotions expressed by the brazilian public. *Comunicar: Media Education Research Journal*, 31(76):117–128. DOI: https://doi.org/10.3916/C76-2023-10.

Fernandes-de Oliveira, G., Massarani, L., Oliveira, T., Scalfi, G., and dos Santos-Jr, M. A. (2023b). The covid-19 vaccine on tiktok: A study of emotional expression in the brazilian contexto. *Evolutionary studies in imaginative culture*, pages 28–45. DOI: https://doi.org/10.56801/esic.v7.i2.3.

Fernandes-de Oliveira, G., Massarani, L., Oliveira, T., Scalfi, G., and dos Santos-Jr, M. A. (2023c). The vaccine on instagram: study of emotions expressed in the brazilian context. *Mediterranean Journal of Communication*, 14(2):283–298. DOI: https://doi.org/10.14198/MEDCOM.2ER4190.

Ferrada, F. and Camarinha-Matos, L. M. (2024). Emotions in human-ai collaboration. In *Working Conference on Virtual Enterprises*, pages 101–117. Springer. DOI: https://doi.org/10.1007/978-3-031-71739-0_7.

Formolo, D. and Bosse, T. (2017). Human vs. computer performance in voice-based recognition of interpersonal stance. In *Human-Computer Interaction. User Interface Design, Development and Multimodality*, pages 672–686. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-58071-5_51.

Fortuna, P., da Silva, J. R., Soler-Company, J., Wanner, L., and Nunes, S. (2019). A hierarchically-labeled portuguese hate speech dataset. In *Proceedings of the Third Workshop on Abusive Language Online*, pages 94–104. DOI: https://doi.org/10.18653/v1/W19-3510.

Fortuna, P. and Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys (CSUR)*, 51(4):1–30. DOI: https://doi.org/10.1145/3232676.

Gao, Q., Ning, H., and Du, B. (2021). A survey on the development of intelligent robots in speech emotion recognition. In *2021 International Wireless Communications and Mobile Computing (IWCMC)*, pages 951–956. DOI: https://doi.org/10.1109/IWCMC51323.2021.9498858.

Gomes, V., Mascarenhas, A., Rodowanski, I., Campos, J., Souza, J., Filho, J. S., and Simões, M. (2024). Using speech and text in emotions recognition. In *Anais do XVI Simpósio Brasileiro de Robótica e XV Workshop de Robótica na Educação*, pages 31–36, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.1109/SBR/WRE63066.2024.10838143.

Gonçalves, V. P., Giancristofaro, G. T., Filho, G. P., Johnson, T., Carvalho, V., Pessin, G., Neris, V. P., and Ueyama, J. (2017). Assessing users' emotion at interaction time: a multimodal approach with multiple sensors. *Soft Computing*, 21:5309–5323. DOI: https://doi.org/10.1007/s00500-016-2115-0.

Goulart, C., Valadão, C., Delisle-Rodriguez, D., Caldeira, E., and Bastos, T. (2019a). Emotion analysis in children through facial emissivity of infrared thermal imaging. *PLoS ONE*, 14(3):e0212928. DOI: https://doi.org/10.1371/journal.pone.0212928.

Goulart, C., Valadão, C., Delisle-Rodriguez, D., Funayama, D., Favarato, A., Baldo, G., Binotte, V., Caldeira, E., and Bastos-Filho, T. (2019b). Visual and thermal image processing for facial specific landmark detection to infer emotions in a child-robot interaction. *Sensors*, 19:2844. DOI: https://doi.org/10.3390/s19132844.

Guimarães, P. R. B. (2014). *Construção e testagem de um instrumento de reconhecimento de expressões faciais emocionais*. PhD thesis, Dissertação de Mestrado em Psicologia Cognitiva, Universidade Federal de Pernambuco.

Halperin, E. (2008). Group-based hatred in intractable conflict in israel. *Journal of Conflict resolution*, 52(5):713–736. DOI: https://doi.org/10.1177/0022002708314665.

Halperin, E. (2011). Emotional barriers to peace: Emotions and public opinion of jewish israelis about the peace process in the middle east. *Peace and Conflict*, 17(1):22–45. DOI: https://doi.org/10.1080/10781919.2010.487862.

Halperin, E., Canetti, D., and Kimhi, S. (2012). In love with hatred: Rethinking the role hatred plays in shaping political behavior. *Journal of Applied Social Psychology*, 42(9):2231–2256. DOI: https://doi.org/10.1111/j.1559-1816.2012.00938.x.

Hammes, L. O. A. and de Freitas, L. A. (2021). Utilizando bertimbau para a classificação de emoções em português. In *Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL)*, pages 56–63. SBC. DOI: https://doi.org/10.5753/stil.2021.17784.

Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. DOI: https://doi.org/10.1145/1014052.1014073.

Huang, M. H. and Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 49:30–50. DOI: https://doi.org/10.1007/s11747-020-00749-9.

Jackson, P. and Haq, S. (2014). *Surrey audio-visual expressed emotion (savee) database*. University of Surrey, Guildford, UK.

Kingeski, R. (2019). Desenvolvimento de um banco de dados de voz com emoções em idioma português brasileiro. Master's thesis, Center of Technological Sciences, Program in Electrical Engineering, State University of Santa Catarina. Joinville.

Leite, J. A., Silva, D., Bontcheva, K., and Scarton, C. (2020). Toxic language detection in social media for brazilian portuguese: New dataset and multilingual analysis. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pages 914–924. DOI: https://doi.org/10.18653/v1/2020.aacl-main.91.

Liang, P. P., Salakhutdinov, R., and Morency, L. P. (2018).

Computational modeling of human multimodal language: The mosei dataset and interpretable dynamic fusion. In *First Workshop and Grand Challenge on Computational Modeling of Human Multimodal Language, vol. 113*, pages 116–125.

Lima, G., Oliveira, A., Silva, F., Pinheiro, L., Luz, E., and Freitas, L. (2024). Desafios sobre a detecção de discurso de Ódio em português brasileiro: Construção de um novo corpus e reflexões sobre o processo. In *Thirtieth Americas Conference on Information Systems, 2024, Salt Lake City*.

Lima, H. M. O. (2021). Banco de vozes brasileiro nas variações das emoções (emovox-br): elaboração e validação. Master's thesis, Department of Phonoaudiology, Federal University of Paraíba, João Pessoa.

Marcacini, R. M., Candido Junior, A., and Casanova, E. (2022). Overview of the automatic speech recognition for spontaneous and prepared speech & speech emotion recognition in portuguese (se&r) shared-tasks at propor 2022. In *CEUR Workshop Proceedings*.

Marques, T. (2023). The expression of hate in hate speech. *Journal of Applied Philosophy*, 40(5):769–787. DOI: https://doi.org/10.1111/japp.12608.

Martinazzo, B. (2010). Um método de identificação de emoções em textos curtos para o português do brasil. Master's thesis, Informatics, Pontifical Catholic University of Paraná (PUCPR), Curitiba.

Martinazzo, B., Dosciatti, M. M., and Paraiso, E. C. (2011). Identifying emotions in short texts for brazilian portuguese. In *IV International Workshop on Web and Text Intelligence (WTI 2012), pp. 16, 2011*.

Martinazzo, B. and Paraiso, E. C. (2010). Identificação de emoções em notícias curtas. In *CLEI-Conferência Latino-Americana de Informática*, volume 1, pages 1–10.

Martínez, C. A., Van Prooijen, J.-W., and Van Lange, P. A. (2022). Hate: Toward understanding its distinctive features across interpersonal and intergroup targets. *Emotion*, 22(1):46. DOI: https://doi.org/10.1037/emo0001056.

Martins, R. A. D. C. (2011). *Projeto, gravação e edição de base de voz para aplicações em síntese e reconhecimento da fala*. PhD thesis, Graduation Department of Electronics and Computations, Polytechnic School, Federal University of Rio de Janeiro.

Matias, J. C., Muller, T. R., Canal, F. Z., and Scotton, G. G. (2021). Ar de junior sa. In Pozzebon, E. and Sobieranski, A. C., editors, *MIGMA*, pages 153–162. CIARP, LNCS 12702, Springer Nature Switzerland AG, JMRS Tavares et al. (Eds.). DOI: https://doi.org/10.1007/978-3-030-93420-0_15.

Maziero, E. G., Pardo, T. A., Di Felippo, A., and Dias-da Silva, B. C. (2008). A base de dados lexical ea interface web do tep 2.0: thesaurus eletrônico para o português do brasil. In *Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web*, pages 390–392. DOI: https://doi.org/10.1145/1809980.1810076.

Mendes, C., Pereira, R., Frazao, L., Ribeiro, J. C., Rodrigues, N., Costa, N., Barroso, J., and Pereira, A. M. (2024a). Emotionally intelligent customizable conversational agent for elderly care: Development and impact of chatto. In *Proceedings of the 11th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*, pages 208–214. DOI: https://doi.org/10.1145/3696593.3696619.

Mendes, R. N., Tavares, S. K., Campos, L. N. M., and Araújo, F. P. (2024b). Mineração de emoções multirrótulo em textos curtos. In *Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL)*, pages 445–450. SBC. DOI: https://doi.org/10.5753/stil.2024.245174.

Morais, D., Maximiano-Barreto, M., Luchesi, B. M., Muniz, M., and Chagas, M. H. N. (2024). Development and validation of a face database for the recognition of facial expressions of basic emotions in the brazilian population. *Psicologia: Reflexão e Crítica*, 37(1):47. DOI: https://doi.org/10.1186/s41155-024-00325-y.

Nascimento, G., Carvalho, F., Cunha, A. M. d., Viana, C. R., and Guedes, G. P. (2019). Hate speech detection using brazilian imageboards. In *Proceedings of the 25th Brazilian Symposium on Multimedia and the Web*, pages 325–328. DOI: https://doi.org/10.1145/3323503.3360619.

Nascimento, G., Duarte, F., and Guedes, G. P. (2018). Emoções em português do brasil: um conjunto de dados e resultados de base. In *Anais do VII Brazilian Workshop on Social Network Analysis and Mining, SBC*. DOI: https://doi.org/10.5753/brasnam.2018.3595.

Nascimento, I. N. d. A. (2016). Elaboração e validação de um banco de expressões faciais de bebês. Master's thesis, Department of Cognitive and Behavioral Neuroscience, Federal University of Paraíba, João Pessoa.

Negrão, J. G. (2019). *Desenvolvimento de um banco de fotografias e vídeos de expressões emocionais de crianças brasileiras de quatro a seis anos. Facial Affective Brazilian Set - children image and vídeo database (FABS)*. PhD thesis, Mackenzie Presbiterian University, São Paulo.

Negrão, J. G., Osorio, A. A. C., Siciliano, R. F., Lederman, V. R. G., Kozasa, E. H., D'Antino, M. E. F., Tamborim, A., Santos, V., de Leucas, D. L. B., Camargo, P. S., Mograbi, D. C., Mecca, T. P., and Schwartzman, J. S. (2021). The child emotion facial expression set: A database for emotion recognition in children. *Front Psychol*, 29(12):666245. DOI: https://doi.org/10.3389/fpsyg.2021.666245.

Novello, B., Renner, A., Maurer, G., Musse, S., and Arteche, A. (2018). Development of the youth emotion picture set. *Perception*, 47(10–11):1029–1042. DOI: https://doi.org/10.1177/0301006618797226.

Oliveira, F., Reis, V., and Ebecken, N. (2023). Tupy-e: detecting hate speech in brazilian portuguese social media with a novel dataset and comprehensive analysis of models. arXiv :2312.17704.

Oliveira, L. and Thomaz, C. (2006). Captura e alinhamento de imagens: Um banco de faces brasileiro. Undergraduate technical report, Department of Electrical Engineering, FEI, São Bernardo do Campo, São Paulo, Brazil.

Oliveira, M. J., Almeida, d. A., Almeida, d. R., and Silva, E. (2014). Speech rate in the expression of anger: A study with spontaneous speech material. In *Speech Prosody*, pages 164–168, In Proc. 7th International Conference on Speech Prosody. DOI: https://doi.org/10.21437/SpeechProsody.2014-21.

Ortony, A., Clore, G., and Collins, A. (1988). *Cognitive Structure of Emotions*. Cambridge University Press.

Pascual-Leone, A., Gilles, P., Singh, T., and Andreescu, C. A. (2013). Problem anger in psychotherapy: An emotion-focused perspective on hate, rage, and rejecting anger. *Journal of Contemporary Psychotherapy*, 43(2):83–92. DOI: https://doi.org/10.1007/s10879-012-9214-8.

Pasqualotti, P. R. and Vieira, R. (2008). Wordnetaffectbr: uma base lexical de palavras de emoções para a língua portuguesa. *RENOTE*, 6(1). DOI: https://doi.org/10.22456/1679-1916.14693.

Pennebaker, J. W., Francis, M. E., Booth, R. J., *et al*. (2001). Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71.

Penteado, C., Ferreira, M. A. S., Pereira, M. A. G., and Chaves, J. M. S. (2022). # vacinar ou não, eis a questão! as emoções na disputa discursiva sobre a aprovação das vacinas contra a covid-19 no twitter. *Política & Sociedade*. DOI: https://doi.org/10.5007/2175-7984.2021.85145.

Pereira, D. A. (2021). A survey of sentiment analysis in the portuguese language. *Artificial Intelligence Review*, 54:1087–1115. DOI: https://doi.org/10.1007/s10462-020-09870-1.

Poletto, F., Basile, V., Sanguinetti, M., Bosco, C., and Patti, V. (2021). Resources and benchmark corpora for hate speech detection: a systematic review. *Lang Resources & Evaluation*, 55:477–523. DOI: https://doi.org/10.1007/s10579-020-09502-8.

Poria, S., Cambria, E., Bajpai, R., and Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information fusion*, 37:98–125. DOI: https://doi.org/10.1016/j.inffus.2017.02.003.

Raso, T. and Mello, H. (2012). The c-oral-brasil i: reference corpus for informal spoken brazilian portuguese. In *Computational Processing of the Portuguese Language*, pages 362–367, Heidelberg. Springer. DOI: https://doi.org/10.1007/978-3-642-28885-2_40.

Rodrigues, P. L., Moraes, R. d. O., Watanuki, H. M., and Hilster, D. d. (2023). Emotions detection in social media posts. *ENEGEP 2023-A contribuição da engenharia de produção para desenvolvimento sustentável das organizações: cadeias circulares, sustentabilidade e tecnologias*. DOI: https://doi.org/10.14488/ENEGEP2023_TN_WG_404_1987_45438.

Romani-Sponchiado, A., Sanvicente-Vieira, B., Mottin, C., Hertzog-Fonini, D., and Arteche, A. (2015). Child emotions picture set (ceps): development of a database of children's emotional expressions. *Psychology & Neuroscience*, 8(4):467–478. DOI: https://doi.org/10.1037/h0101430.

Rosa, J. J. (2017). Reconhecimento automático de emoções através da voz. Graduation dissertation, Information Systems, Federal University of Santa Catarina, Florianópolis.

Rossettini, G., Conti, C., Suardelli, M., Geri, T., Palese, A., Turolla, A., Lovato, A., Gianola, S., and DellIsola, A. (2021). Covid-19 and health care leaders: how could emotional intelligence be a helpful resource during a pandemic? *Physical Therapy*, 101(9):pzab143. DOI: https://doi.org/10.1093/ptj/pzab143.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161. DOI: https://doi.org/10.1037/h0077714.

Saganowski, S. (2022). Bringing emotion recognition out of the lab into real life: Recent advances in sensors and machine learning. *Electronics*, 11(3):496. DOI: https://doi.org/10.3390/electronics11030496.

Santos, A. G. L., Becker, K., and Moreira, V. P. (2014a). Mineração de emoções em textos multilíngues usando um corpus paralelo. *29th SBBD proceedings*, pages 127–136. ISSN 2316-5170.

Santos, A. G. L., Becker, K., and Moreira, V. P. (2014b). Um estudo de caso de mineração de emoções em textos multilíngues. In *Anais do III Brazilian Workshop on Social Network Analysis and Mining*, pages 140–151.

Scherer, K. (2005). What are emotions? and how can they be measured? *Social Science Information*, 44:695–729. DOI: https://doi.org/10.1177/0539018405058216.

Scherer, K. R., Banse, R., and Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-cultural psychology*, 32(1):76–92. DOI: https://doi.org/10.1177/0022022101032001009.

Schuller, B. W. (2018). Speech emotion recognition: Two decades in a nutshell, benchmarks, and ongoing trends. *Communications of the ACM*, 61(5):90–99. DOI: https://doi.org/10.1145/3129340.

Schuller, D. and Schuller, B. W. (2018). The age of artificial emotional intelligence. *Computer*, 51(9):38–46. DOI: https://doi.org/10.1109/MC.2018.3620963.

Silva, R., Valter, M. F., and Souza, M. (2020). Interaffection of multiple datasets with neural networks in speech emotion recognition. In *Anais do XVII Encontro Nacional de Inteligência Artificial e Computacional*, pages 342–353. DOI: https://doi.org/10.5753/eniac.2020.12141.

Silva, R. O., Losada, J. C., and Borondo, J. (2023). Analyzing the emotions that news agencies express towards candidates during electoral campaigns: 2018 brazilian presidential election as a case of study. *Social Sciences*, 12:458. DOI: https://doi.org/10.3390/socsci12080458.

Silva, W. d., Barbosa, P. A., and Abelin, Å. (2016). Cross-cultural and cross-linguistic perception of authentic emotions through speech: An acoustic-phonetic study with brazilian and swedish listeners. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 32:449–480. DOI: https://doi.org/10.1590/0102-445003263701432483.

Siqueira, E. S., Fleury, M. C., Lamar, M. V., Drachen, A., Castanho, C. D., and Jacobi, R. P. (2023). An automated approach to estimate player experience in game events from psychophysiological data. *Multimedia Tools and Applications*, 82:19189–19220. DOI: https://doi.org/10.1007/s11042-022-13845-5.

Siqueira, E. S., Santos, T. A., Castanho, C. D., and Jacobi, R. P. (2018). Estimating player experience from arousal and valence using psychophysiological signals. In *17th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, pages 107–10709. IEEE. DOI:

https://doi.org/10.1109/SBGAMES.2018.00022.

Siqueira, V., Lunardi, M. G., Silva, W., and R. L. H. Costa, a. S. S. T. (2024a). A dataset of polarities and emotions from brazilian portuguese play store reviews. DOI: https://doi.org/10.5281/zenodo.10823148.

Siqueira, V. X., Costa, R. L. H., Soares, T. S., Lunardi, G. M., and Silva, W. (2024b). Dataset anotado de sentimentos a partir de comentários de aplicativos móveis. In *Dataset Showcase Workshop (DSW)*, pages 65–76. SBC. DOI: https://doi.org/10.5753/dsw.2024.243926.

Sneddon, I., McRorie, M., McKeown, G., and Hanratty, J. (2012). The belfast induced natural emotion database. *IEEE Trans. Affect. Comput.*, 3:32–41. DOI: https://doi.org/10.1109/T-AFFC.2011.26.

Souza, F., Nogueira, R., and Lotufo, R. (2020). Bertimbau: pretrained bert models for brazilian portuguese. In *Brazilian conference on intelligent systems*, pages 403–417. Springer. DOI: https://doi.org/10.1007/978-3-030-61377-8_28.

Souza, J. M. S., Alves, C. D. S. M., Cerqueira, J. D. J. F., Oliveira, W. L. A. D., Pires, O. M., Santos, N. S. B. D., Wyzykowski, A. B. V., Pinheiro, O. R., Almeida, F. D. G. D., da Silva, M. O., and Barbosa, J. D. V. (2024). Facial biosignals time–series dataset (fbiot): A visual–temporal facial expression recognition (vt-fer) approach. *Electronics*, 13(24):4867. DOI: https://doi.org/10.3390/electronics13244867.

Souza, M., Vieira, R., Busetti, D., Chishman, R., and Alves, I. M. (2011). Construction of a portuguese opinion lexicon from multiple resources. In *Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology*, pages 59–66.

Souza, V. F. (2022). Aplicação de um minerador de emoções em fóruns de discussão de um massive open online course (mooc) brasileiro: uma abordagem utilizando o algoritmo naive Bayes. *EaD em Foco*, 12(2):e1732. DOI: https://doi.org/10.18264/eadf.v12i2.1732.

Spezialetti, M., Placidi, G., and Rossi, S. (2020). Emotion recognition for human-robot interaction: Recent advances and future perspectives. *Frontiers in Robotics and AI*, page 145. DOI: https://doi.org/10.3389/frobt.2020.532279.

Steiner-Correa, F., Viedma-del, J. M. I., and Lopez-Herrera, A. G. (2018). A survey of multilingual humantagged short message datasets for sentiment analysis tasks. *Soft Comput*, 22:8227–8242. DOI: https://doi.org/10.1007/s00500-017-2766-5.

Sturgill, R., Martinasek, M., Schmidt, T., and Goyal, R. (2021). A novel artificial intelligence-powered emotional intelligence and mindfulness app (ajivar) for the college student population during the covid-19 pandemic: quantitative questionnaire study. *JMIR Formative Research*, 5(1):e25372. DOI: https://doi.org/10.2196/25372.

Tejada, J., Freitag, R. M. K., Pinheiro, B. F. M., Cardoso, P. B., Souza, V. R. A., and Silva, L. S. (2022). Building and validation of a set of facial expression images to detect emotions: a transcultural study. *Psychological Research*, 86(6):1996–2006. DOI: https://doi.org/10.1007/s00426-021-01605-3.

Thomaz, C. E. and Giraldi, G. A. (2010). A new ranking method for principal components analysis and its application to face image analysis. *Image Vis. Comput*, 28:902–913. DOI: https://doi.org/10.1016/j.imavis.2009.11.005.

Torres, N. J. R., Mano, L. Y., and Ueyama, J. (2018). Verbo: voice emotion recognition database in portuguese language. *Journal of Computer Science*, 14(11):1420–1430. DOI: https://doi.org/10.3844/jcssp.2018.1420.1430.

Trajano, D., Bordini, R., and Vieira, R. (2024). Olid-br: Offensive language identi¦cation dataset for brazilian portuguese. *Language Resources and Evaluation*, 58(4):1263–1289. DOI: https://doi.org/10.1007/s10579-023-09657-0.

Vargas, F., Carvalho, I., de Góes, F. R., Pardo, T., and Benevenuto, F. (2022). Hatebr: A large expert annotated corpus of brazilian instagram comments for offensive language and hate speech detection. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference, Association for Computational Linguistics*, pages 7174–7183. DOI: https://doi.org/10.63317/544cc4es6d8a.

Vieira, L. C. (2017). *Assessment of fun from the analysis of facial images*. PhD thesis, Institute of Mathematics and Statistics, University of Sao Paulo.

Vogt, T., André, E., and Wagner, J. (2008). Automatic recognition of emotions from speech: a review of the literature and recommendations for practical realisation. *Affect and Emotion in Human-Computer Interaction: From Theory to Applications*, pages 75–91. DOI: https://doi.org/10.1007/978-3-540-85099-1_7.

Wang, J. Z., Zhao, S., Wu, C., Adams, R. B., Newman, M. G., Shafir, T., and Tsachor, R. (2023). Unlocking the emotional world of visual media: An overview of the science, research, and impact of understanding emotion. In *Proceedings of the IEEE, 2023*. DOI: https://doi.org/10.1109/JPROC.2023.3273517.

Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., Gao, S., Sun, Y., Ge, W., Zhang, W., and Zhang, W. (2022). *A systematic review on affective computing: Emotion models, databases, and recent advances*. Information Fusion.

Weitzel, L., Daroz, T. H., Cunha, L. P., Helde, R. V., and de Morais, L. M. (2023). Investigating deep learning approaches for hate speech detection in social media: Portuguese-br tweets. In *2023 18th Iberian Conference on Information Systems and Technologies (CISTI), IEEE*, pages 1–5. DOI: https://doi.org/10.23919/CISTI58278.2023.10211751.

Wierzbicka, A. (1999). *Emotions across languages and cultures: Diversity and universals*. university press. Cambridge University Press, Cambridge.

Wingenbach, T. S. H., Morello, L. Y., Hack, A. L., and Boggio, P. S. (2019). Development and validation of verbal emotion vignettes in portuguese, english, and German. *Frontiers in Psychology*, 10:1135. DOI: https://doi.org/10.3389/fpsyg.2019.01135.

Zampar, S. (2010). O uso do grotesco no humor das rádios fm de público jovem de são paulo - rádio mix e 89 em. Master's thesis, Institute of Social Sciences and Communication, University of Paulista, São Paulo.