Lexicography Saves Lives (LSL): Automatically Translating Suicide-Related Language

Annika Marie Schoene[1], John E. Ortega [1], Rodolfo Joel Zevallos[2], Laura Haaber Ihle [1],
[1]Northeastern Universty, Institute for Experiential AI, [2]Barcelona Supercomputing Center,
Correspondence: amschoene@gmail.com

Abstract

Recent years have seen a marked increase in research that aims to identify or predict risk, intention or ideation of suicide. The majority of new tasks, datasets, language models and other resources focus on *English* and on suicide in the context of Western culture. However, suicide is global issue and reducing suicide rate by 2030 is one of the key goals of the UN's Sustainable Development Goals[1]. Previous work has used *English* dictionaries related to suicide to translate into different target languages due to lack of other available resources. Naturally, this leads to a variety of ethical tensions (e.g.: *linguistic misrepresentation*), where discourse around suicide is not present in a particular culture or country. In this work, we introduce the *'Lexicography Saves Lives Project'* to address this issue and make *three* distinct contributions. First, we outline ethical consideration and provide overview guidelines to mitigate harm in developing suicide-related resources. Next, we translate an existing dictionary related to suicidal ideation into 200 different languages and conduct human evaluations on a subset of translated dictionaries. Finally, we introduce a public website to make our resources available and enable community participation.

1 Introduction

Each year more than 700,000 people die by suicide worldwide (Organization et al., 2021), where for each suicide there are many more attempts[2] and often numbers are underestimated due to underreporting or misclassification (Snowdon and Choi, 2020). There are multiple factors at play that contribute, which include but are not limited to (i) social stigma, (ii) cultural and/or (iii) legal concerns (Dattani et al., 2023). There are a variety of efforts that focus on developing new prevention, screening and risk identification strategies to reduce suicide rates not only from the medical, public health and policy community (Morrow et al., 2022; Little et al., 2016; Denneson et al., 2016), but also from Machine Learning (ML) and Natural Language Processing (NLP) community (Kim et al., 2023; Badal and Depp, 2022; McCoy et al., 2016). One area of this work focuses on detecting suicidal risk, intent or ideation from social media (Coppersmith et al., 2015; Du et al., 2018; Schoene et al., 2023) by utilizing keyword detection approaches based on dictionaries (Sinha et al., 2019; Ji et al., 2020). However, the vast majority of this work are developed in *English* and for a western culture, where fewer resources exist in other languages and close to none in low-resource languages. This lack of resources has spurred efforts to automatically translate existing dictionary resources from *English* and *Chinese* to *Korean* (Lee et al., 2020) to predict suicidality. Despite some success of such approaches, there are still a number of challenges that limit the usefulness where automatic translations often don't take cultural context or linguistic differences into account (Ortega and Church, 2023). For example, when translating from *English* to *German* the phrase *'my suicide letter'* is translated into *'Mein Selbstmordbrief'*. Whilst this is a grammatical correct translation, the phrase itself would be rarely if at all used in German and a more commonly used phrase would be *'Mein Abschiedsbrief'*. Similarly, there may be words and phrases in a target language that do not exist in the source language and therefore would be missing in the final translated dictionary. This also raises a set of linguistic and ethical concerns, such as the consequences of designing automatic suicide ideation detection resources/tools and their usefulness in practice.

Despite these challenges, the potential use of machine translation in the detection of suicide cases is a field that deserves further investigation and careful approaches, as it could offer an additional

[1]https://sdgs.un.org/goals/goal3
[2]https://www.who.int/news-room/fact-sheets/detail/suicide

Proceedings of the 31st International Conference on Computational Linguistics, pages 3179–3192
January 19–24, 2025.

means of support for suicide prevention on a global scale, provided that the mentioned challenges are appropriately addressed. To grow and improve resource and practices around the development of resources that aid suicide ideation detection in different languages we introduce the *'Lexicography Saves Lives'* and make the following contributions:

- First, we describe general ethical issues as well as ethical concerns related to linguistic misrepresentation that arise with automatic translation in the context of developing resources for suicide ideation detection. Then we provide an overview set of ethical guidelines and mitigation strategies that can aid in developing, designing, and implementing suicide-related resources in the future (section **??**).

- Next, we translate a seed dictionary proposed by O'dea et al. (2015) into 200 different languages (Team et al., 2022a) using the Flores 101 evaluation benchmark (Goyal et al., 2021) (section 3). We then conduct a human evaluation on a subset of *5* translated dictionaries to better understand the quality. We look at *7* different variables, including *Adequacy*, *Fluency*, *Spelling Errors*, *Culture*, *Context*, *Alternative Translations* and *Contributions in local language*. For *5* out of the *7* variables we quantify our findings to give a score that indicates the quality of the translated dictionary (section 4).

- Finally, we have created a public website to make our translated and evaluated resources available, provide constant progress updates on what the status of each dictionary is in it's respective language and enable community participation (section 5).

Lexicography Saves Lives Project: Beyond the scope of this paper, we hope that with this project we raise awareness around the development and deployment of old and new resources (e.g.: datasets, lexicons etc.), algorithms and ultimately tools designed for Mental Health, and especially suicide prevention. We very much see this as a starting point of a wider conversation around how as a community we need to advocate for (i) transparency around harms and benefits of doing such work, (ii) true interdisciplinary research, and (iii) community involvement.

2 Related Work

Detecting suicide-related language Detection methods for suicidal intent, ideation or risk based on machine learning have evolved significantly over the past decades, and various techniques have been employed to enhance model accuracy. Traditionally, feature engineering has been a crucial component of these methods, where features extracted from text using dictionaries play a pivotal role in training machine learning models Sarsam et al. (2021); Birjali et al. (2017); Abboute et al. (2014); Okhapkina et al. (2017); Ji et al. (2022). Our efforts are closely related to previous work that investigated methods and resources for suicide ideation detection. Collecting annotated data for mental health related tasks is notoriously difficult and suicidal ideation is no exception. Work in this area usually relies on self-reports (Coppersmith et al., 2015), heuristics based on the presence of specific keywords (Du et al., 2018) and phrases or words (Burnap et al., 2017; Coppersmith et al., 2015) in lexicons Sawhney et al. (2018).

More specifically, lexicons have typically been developed in collaboration with or by domain experts (Gaur et al., 2019), but more recently computational methods have been used generate new lexicons using social media data (Lee et al., 2022; Lv et al., 2015). However, the vast majority of research focuses on developing new datasets, resources (e.g.: language models, lexicons) and tasks only in *English* (Du et al., 2018; Schoene et al., 2023; Mishra et al., 2019; Sawhney et al., 2021; Cao et al., 2019). Some existing work in languages other than *English* often focus on high-resource languages such as *Spanish* (Valeriano et al., 2020; Ramírez-Cifuentes et al., 2020), Arabic (Hassib et al., 2022; Benlaaraj et al., 2022) or *Chinese* (Huang et al., 2014; Lv et al., 2015) or and very few investigate low-resource languages (e.g.: *Filipino* or *Taglish* (Astoveza et al., 2018)).

Other work (Moslem et al., 2023) focused on the use of adaptive machine translation to create translations for domain-specific text in low-resource languages like *Kinyarwanda*. They showed that generation tools such as GPT-3 (Brown et al., 2020; Ouyang et al., 2022) and Bloom (Davis, 2023) were inadequate when for domain-specific text in low-resource situations.

Lastly, Lewis et al. (2011) provided a cookbook for MT in crisis situations. Their work could be a good guide to consider when translating texts such as suicide or other life-threatening texts because it outlines steps of actions to take during a crisis. While there is no work that completely covers what to do for translation purposes in suicidal situations in a low-resource community, we appreciate their cookbook and suggest that more work be done like this for suicidal texts.

Ethical considerations for Mental Health and Suicide Ethics are another critical concern when considering the implementation of suicide detection systems based on machine translation. People's privacy must be rigorously protected, and any approach in this regard should take into account the ethical implications of online surveillance, the collection of sensitive data, and impact of peoples health. In the past, a few efforts have been made to raise awareness around ethical tensions and issues around online suicide prevention tools, methods and approaches (Orr et al., 2022). The majority of existing work in this space has investigated and reviewed existing apps, products or platforms that are already in use (Gomes de Andrade et al., 2018; Martinengo et al., 2019; Larsen et al., 2015; Jha et al., 2023; Braciszewski, 2021) and much of existing recommendations have been based on practical experience rather than foundational work in bioethics or AI ethics. To the best of our knowledge there has been no effort focused on settings where MT systems are used to automatically translate suicide-related language and has taken into considerations linguistic and cultural factors.

subsectionLinguistic Imbalance and Misrepresentation Ethical concerns arise from the imbalance between various languages and language groups or from underlying, but often unarticulated, assumptions about what language is and how it functions.

Imbalances There is an inherit imbalance between high-resource languages and low-resource languages (Ortega and Church, 2023). For example, some languages are far better represented in data sources due to the number of speakers of said language or tech imperialism (Kwet, 2019), and therefore provide a much stronger source for training. This ultimately results in better functioning tools within those language groups. In contrast, low-resource languages are often limited by the size of the data sources, which can result in lower functionality of the finished tool for those language groups. An unintended consequence of this problem is that high-resource language speakers will be better served by the tools developed, while communities who belong to the low-resource language speakers will be under-served, resulting in an obvious distributive injustice, especially if speakers of low-resource languages will still be subjected to the tools, but the tools will be less efficient on such language speakers.

Misrepresentation poses a risk of low functionality in suicide ideation detection tools in low-resource languages. This can occur when a high-resource language forms the only basis of inquiry, against which all other translations are made. For example, one can imagine a simplistic approach to multilingual suicide ideation detection, where *English* is used as a starting point to identify words related to death and suicide. The list of words is then translated into a variety of low-resource languages, and the results are used to form the basis of the automated suicide ideation detection tools without human intervention. The linguistic misrepresentation will follow from an underlying assumption about what language is and how it functions. Namely that words in language refer to objects or strictly defined concepts in the world and that the meaning of each word follows directly from the object/concept it refers to. If that was the case, it would indeed be possible to simply translate the list from one language to another and obtain a well functioning output. However, a distinction made in philosophy of language may assist in establishing why this understanding is problematic. This distinction is between meaning and reference (Putnam, 1981), where the reference of a word to refer to the object it points to, but the meaning of the word to refer to the intention behind the word (Speaks, 2010; Frege, 1892; Wittgenstein, 2010; Russell, 1905).

Therefore, if multilingual datasets are developed merely by translating words from a high-resource language to a range of low-resource languages, with no other additional initiatives, the result will be a list of words that reference death and suicide. However, it will not encompass the linguistic meaning of suicide and death in each of the languages to which the translation is made. This becomes evident when one considers how many of the common expressions concerning death or suicide that are perfectly meaningful to language users, but do not contain any reference to death or suicide at all; the *English* phrase *'to kick the bucket'*, has the Danish version *'at stille træskoene" (to put down the wooden clogs)'*, and the Turkish *'nallari dikmek"* *(to put up the horseshoes)'*. Without a thorough understanding of the fact that a large part of the linguistic meaning we establish when talking about suicide or death are purely metaphorical, allegorical, context dependent, and deeply local to language users, suicide ideation tools will never catch the myriads of ways we meaningfully talk about those topics, and therefore never be as efficient as they could be.

The ethical effect of this linguistic misrepresentation - a type of representational injustice -is also distributive injustice (Zalta et al., 1995), because the tool will have lower precision and functionality in low-resource languages, where attention to how death and suicide are encompassed in language has merely been replaced with a checklist of the words that correspond to those in the high resource language used as the baseline.

2.0.1 Mitigation strategies

There are various ways to mitigate the aforementioned risks, which include but are not limited to (i) paying close attention to the quality and size of data sources, securing as balanced an approach as possible and (ii) letting each language speak for itself in its respective cultural contexts. It is critical, that local speakers of all languages, but especially low-resource languages are actively included in the translation process, not merely to check if the automated translations make sense, but to add words and phrases from their own language that relate to death and suicide, and to delete those automatically generated translations that are not meaningful in their language. The importance of integrating a strong cultural context becomes clear, when one understands that meaning in language arises through the exact, specific context of each language: To put down the wooden clogs, stems from a time when all Danes were farmers and only had one pair of shoes and putting them down meant one was dead (Tangherlini, 2013). Meaning in language does not merely stem from their reference to objects in the world (Wittgenstein, 2010), and for that reason only local language speakers have the authority to define language about death and suicide in their language group. An under-representation of language is an under-representation of the cultural context, in which that language exists. It is evident that even within a particular language group there are linguistic variations and the potential for under-representation of specific sub-groups. For that reason, one should aim for as broad a group of collaborators as possible, within each language group.

2.1 Design, Autonomy and Justice

Design Both the design and implementation of suicide prevention tools demand a thorough focus on ethics and on the potential implications the technologies in question may have on those who are subjected to them. This requires a theoretically founded and methodological approach to ethics, where foundational work in bioethics (Beauchamp et al., 2008; of Health et al., 2016) and AI ethics (Floridi et al., 2021; Canca, 2020) can provide valuable insights. Even when tools are designed with good intention and the aim to help and protect the most vulnerable, they also have the potential to cause harm or wrongdoing to that same group. This should always be acknowledged by those with the power to develop and implement such tools, and all measures should be taken to ensure that no harm comes to those who are most vulnerable to it. Even if this leads to not implementing the tool is the least potentially harmful path forward. This is especially important in settings, where tech-based initiatives (e.g.: platforms, apps etc.) do not not fall under the scope of binding regulations or organizationally enforced ethical procedures (Celedonia et al., 2021). In fact, no such guidelines are necessarily developed or implemented. Nor does developing (or implementing) suicide prevention tools demand any professional training in mental health or health care in general. In the worst case, this can result in a setting where expertise is low, ethical concerns are overlooked or missing and the output actively targets the most vulnerable.

Autonomy A clear and available introduction to how the tools function should be available to everyone subjected to them, just as it should be possible to inquire about the basis on which they reached specific conclusions. Individuals should be clearly informed about the use of suicide ideation detection tools and should explicitly consent to being subjected to them. To ensure individual autonomy, and due to the potentially severe stigmatization of subjects and the risk of unjust representation in the case of false positives, the utmost care should be given to designing any tools that aim to detect suicide ideation, intent or risk. Any tools developed to prevent suicide require close monitoring of personal data to function, and thus entails a potentially severe privacy violation. This trade off between saving lives and closely monitoring personal data should be made only if the tools are actually effective. Ensuring and documenting efficiency is therefore crucial, in order to justify the potential privacy violation involved.

Justice It is vital, to ensure that any resources or tools are subjected to ongoing bias and fairness audits, so as to avoid that the potential burdens of the system (false positives) are not disproportionately distributed on particular groups or individuals, but equally distributed across the population. This is especially important in this context, because of the potentially stigmatizing effect of false positives and the emotional harm they may incur. Furthermore, it is essential to ensure that local regulatory and cultural contexts around suicide are taken into consideration. In 25 countries, not only suicide, but suicide attempts are illegal and for example, in the Bahamas, it is punished with life imprisonment (Mishara and Weisstub, 2016). In such a setting, a suicide prevention tool can become an instrument of power, surveilling citizens for signs of criminal behavior. Therefore, the potential positive effects of suicide prevention tools can outweigh the negative effects and careful measures should be taken prior to development.

Overall, it is crucial to establish strong ethics practices, clear procedures for complying with them, and processes for documenting that compliance. As the tools in question develop over time, this requires an ongoing involvement with their ethical implications and a thorough integration of ethics into all stages of the development and deployment process, from research, to design, development and deployment. This paper does not offer such an extensive ethics framework, but rather provides an overview guidance on the ethical risks involved in designing, developing, and deploying suicide prevention tools, and to clarify the scope of potential ethical pitfalls, when it comes to developing new resources.

3 Automatic translation

The original lexicon containing words and phrases related to suicidal ideation was proposed by O'dea et al. (2015) and includes the following 50 words and phrases. We provide the full list of phrases/words that express suicidal ideation based on work by Sawhney et al. (2018) and originally developed by O'dea et al. (2015):

- suicidal, kill myself, my suicide letter, end my life, never wake up, suicide pact, die alone, wanna die, why should I continue living, to take my own life, suicide, can't go on, want to die, be dead, better off without me, better off dead, dont want to be here, go to sleep forever, wanna suicide, take my own life, suicide ideation, not worth living, ready to jump, sleep forever, suicide plan, tired of living, die now, commit suicide, thoughts of suicide, depressed, slit my wrist, cut my wrist, slash my wrist, do not want to be here, want it to be over, want to be dead, nothing to live for, ready to die, not worth living, I wish I were dead, kill me now, hit life, think suicide, wanting to die, suicide times, last day, feel pain point, alternate life, time to go, beautiful suicide, hate life

3.1 Experiments

We translate from English into 200 languages originally proposed by the No Language Left Behind (NLLB) (Team et al., 2022b) evaluation dataset for low-resource languages. The original language list from NLLB can be found in the Appendix (Section **??**). The automated machine translation (MT) system used for translation purposes is based on the Flores 101 evaluation benchmark (Goyal et al., 2021) which was extended to cover 200 languages[3]. We use the Fairseq research toolkit[4] (Ott et al., 2019) with the transformer-based (Vaswani et al., 2017) pre-trained language model (PLM) baseline which is a multi-language model that accepts English as the input and is capable of translating to 200 languages. We aim in our experiments to specifically validate suicidal language in one direction, English–>*target language*. For future work, we plan on creating several automated heuristics to automatically validate the translations of new entries and update the current ones where needed.

4 Multi-lingual Lexicon Evaluation

We evaluate our automatically translated lexicons by using qualitative and quantitative measures to gain a deeper insight into the quality and appropriateness of our translations.

4.1 Human Evaluation

We rely on previous work (Ortega and Church, 2023; Castilho et al., 2018; O'Brien, 2017) that

[3]https://github.com/facebookresearch/flores/blob/main/flores200/README.md

[4]https://github.com/facebookresearch/fairseq

describe some of the major facets and caveats of evaluating translations into target languages, including variables that are important to languages that are scarce. For this, we introduce five different variables to evaluate the qualitative validity of each lexicon and two additional *free text* variables for contributions. Here we list and described each variable and how its measured for each words/phrase in a translated lexicon:

- **Adequacy** Similar to Castilho et al. (2018) we ask annotators to evaluate how adequate a translation is by asking *'How much of the source text meaning has been retained in the translated language?'*. The goal is to rate on a Likert scale (Jebb et al., 2021) from 1 to 4 how much meaning was retained, where 1 corresponds to *no meaning retained* and 4 *all meaning retained*.

- **Fluency** Similar to *Adequacy* we ask annotators how fluent a translation is by asking *How fluent is the translation?*. As before, we employ a Likert scale from 1 to 4, where 1 is *'no fluency'* and 4 is *'native'*. While 'adequacy' would typically measure the degree to which the translation reflects the content of a source sentence, here, "fluency" measures how fluent a translation reads without referring to either the source sentence or reference translation (Graham et al., 2017).

- **Spelling Errors** For each translated word/phrase we ask annotators to score one point if a translation contains errors, such as 'misspelled words', 'missing words', 'added words' or 'incorrect word order'.

- **Cultural Acceptability** Given that suicidal language is not universal (Kirtley et al., 2022), we aim to better facilitate target translations in their cultural context. We propose the following approach: for phrases like "bite the dust" that may not translate well into other languages. If the lexicon's target language translation does not match the source language's intent, the participant is asked to select "no", otherwise when the source language's lexicon matches the target translation, the participant selects "yes". *Question: Does this word/phrase occur in your culture?*

- **Context** After verifying if a word/phrase exists in a cultural context, we also want to verify that it exists in the context of language related to suicide. As previously outlined, if the word/phrase occurs in this cultural context participants can either select "yes' or "no' and have the option to add variation if applicable. *Question: Does this word/phrase occur in the context of suicidal ideation?*

Free text variables:

- **Alternative Translation** In addition to the previous variables, we would also like to take into consideration further feedback. Here we participants can either select 'no' or add comments related to each entry.

- **Contributions in local language** In this section we ask participants to add (i) words related to death, (ii) words related to suicide, (iii) expressions/metaphors related to dying and (iv) expressions/metaphors related to suicide.

4.2 Metrics

Translating sensitive content such as language related to suicide ideation requires both high translation quality and appropriateness for the target culture and context (Kirtley et al., 2022). Quantitative evaluation metrics can complement human assessment to provide a comprehensive analysis of translation adequacy. We use quantitative metrics based on the 5 variables proposed in section 4.1 to measure appropriateness using human judgments. Each metric follows the same formula, however the meaning of each score may differ. Simply put, we take the total number of entries in the dictionary (N) and calculate the arithmetic mean $\overline{x}$ using the sum of submitted evaluations E:

$$\overline{x} = \frac{\sum_E}{N} \tag{1}$$

Furthermore, we calculate each metric per submitted dictionary and then take the average of all submissions per language. For scores related to *Adequacy* and *Fluency* a dictionary can score a maximum value of *4*, meaning all meaning and fluency has been retained in the translation respectively. The best score for *Spelling Errors* is *0*, which shows that no spelling errors were made by the MT system. Next, for *Cultural* and *Contextual acceptability*, the best score is *1* showing that all translations are deemed appropriate.

4.3 Pilot Evaluation

Given the large number of lexicons in a variety of languages, we conduct a round of pilot evaluations of our proposed qualitative measures by inviting native speakers of each language to participate. First, we identify annotators for a select number of languages much as was done in previous work by Facebook (Costa-jussà et al., 2022; Barrault et al., 2023). Each annotator is given (i) the original dictionary alongside the translation and (ii) a codebook with an example evaluation for reference [5]. At this stage, we specifically ask native speakers in the general public, who have no medical, psychological or behavioral health training. For our first round of evaluations we chose *5* languages, where in Table 1 we show the languages and respective scores. Each dictionary was evaluated by at least 2 participants.

4.4 Results

In Table 1 we report the scores for each translated dictionary, where we find that for both *Adequacy* and *Fluency* translations are high overall (4 out of 5 lexicons score over 3.0 in both categories) with *Danish* having the highest and *Finish* the lowest score. Furthermore, there are fewer *Spelling Errors* in both *Danish* and *German*, whereas *Galician* has the highest score. This could be due to both languages having being better resource representation and therefore functionality is higher compared to low-resource languages. For *Cultural* and *Contextual Acceptability* scores are lowest in *Finish* indicating that the proposed translations of suicide-related language are not applicable either culturally or contextually.

Alternative Translations For each translated lexicon, we also asked participants to provide any alternative translations that may exist in addition to the proposed translation. In Table 2 we list a sample of alternative translation for each dictionary, where we find that in some instances more colloquial terms, metaphors or analogies are more appropriate instead of the literal translation provided by the MT system.

Language Contributions In addition to alternative translations, we have asked for contributions in local languages that may not be covered in the originally proposed dictionary. In Table 3 we list examples in *Danish* and *Finish* as we have been given contributions in those languages. As expected, we find that some terminology that is more relevant to the topic in the target language is not represented in the source dictionary. We hope in future iterations of this work we will be able to solicit more contributions to grow resources and make them available to the wider research community.

5 Website for Participation

One of the main aims for this work is to increase community engagement and participation to ensure the responsible development of the proposed lexicons. Therefore we have created a platform [6] with three main objectives: to provide visual summaries of the progress of the project, to disseminate the dictionaries and other resources generated and finally to collect local feedback and new input from users that help us improve results. The visualization problem abstraction framework defined by Munzner, helps defining data visualization problems by addressing what data is being visualized, why is it visualized (tasks), and how is it visualized (visualization idioms, marks and channels). With this platform we aim to address two main high level analysis tasks (why) as defined by the framework: to *present* the current progress of the project and to *produce* feedback data that helps us improve the project. For the *present* task, the website will include interactive visual summaries (Summarization target action from the framework) that will help users understand the current progress of the project and what main areas have been covered. Moreover, the interactive part of this visualization idioms would allow the user to explore the results and find specific insights that are more relevant for their context (Search action tasks). Furthermore, the platform with also help as a feedback collection mechanism (Produce task), which will allow the users to improve the repository and to adjust the data to the specifics characteristics of each culture and language.

6 Conclusion, Limitations and Future Work

In this work, we have introduced the *'Lexicography Saves Lives Project'* and a set of broad ethical considerations and guidelines that look at how this work can be carried forward responsibly. Furthermore, we have automatically translated an existing lexicon containing 50 words and phrases related to

[5] Link to code book: made available upon publication

[6] Made available upon publication: `www.dummy.com`

Language	Adequacy	Fluency	Spelling Errors	Culture	Context
Catalan	3.68	3.6	3.7	0.76	0.8
Danish	3.74	3.6	0.02	0.98	0.86
Finish	2.8	2.58	0.08	0.68	0.48
Galician	3.48	3.41	0.29	0.96	0.96
German	3.12	2.84	0.02	0.84	0.74

Table 1: Evaluation of five translated dictionaries using quantative metrics introduced in section 4.2.

Original word/phrase	Proposed Translation	Alternative Translation	Language
to take my own life	prendre la meva pròpia vida	prendre'm la vida	Catalan
	Otan oman henkeni	otan henkeni	Finish
	para quitar a miña propia vida	para quitarme a vida	Galician
	Mein eigenes Leben zu nehmen	mich umbringen	German
slit my wrist	T'he tallat la mà	M'he tallat la mà	Catalan
	Skær mit håndled	skærer mit håndled over	Danish
	cortoume o pulso	cortar o pulso	Galician
	Schneide mir das Handgelenk ab	ritzen	German
go to sleep forever	Gå i seng for evigt	Sove for evigt	Danish
	Vai durmir para sempre	vai adormecer para sempre	Galician
	Schlafe für immer	für immer schlafen	German

Table 2: Examples of alternative translations contributed by participants.

Language	Language Contribution	English Translation
Danish	kradse af	die
Danish	stille træskoene	die
Danish	forenes med gud	unite with the lord
Danish	springe i døden	commit suicide by jumping
Danish	Dø for egen hånd	Suicide (die by your own hand)
Finish	itsetuho	self-harm
Finish	itsetuhoisuuden tilalla	about to commit self-harm, suicidal

Table 3: Sample of additional language contributions and their translations submitted by participants

suicide into 200 language using Flores 101. Subsequently, we conducted a round of human pilot evaluations on a subset of translated lexicons and introduced a set of metrics to score the quality of a translated lexicon. We described some of our initial findings and outlined our website for participation, which will be launched upon publication of this work. In future iterations of this project we hope to address the following limitations: Firstly, we will develop a comprehensive set of ethics guidelines and assessment strategies to provide ongoing and improved monitoring of this project. Next, we will improve the methodology behind the lexicon quality scores and identify a means of measuring overall quality. As part of this effort, we also plan to compare our current scores against existing and more traditional measures of translation quality in MT systems, such as comparing embedding similarities. Finally, we hope to solicit more evaluations and contributions in local languages.

References

Amayas Abboute, Yasser Boudjeriou, Gilles Entringer, Jérôme Azé, Sandra Bringay, and Pascal Poncelet. 2014. Mining twitter for suicide prevention. In *Natural Language Processing and Information Systems: 19th International Conference on Applications of Natural Language to Information Systems, NLDB 2014, Montpellier, France, June 18-20, 2014. Proceedings 19*, pages 250–253. Springer.

Ghelmar Astoveza, Randolph Jay P Obias, Roi Jed L Palcon, Ramon L Rodriguez, Bernie S Fabito, and Manolito V Octaviano. 2018. Suicidal behavior detection on twitter using neural network. In *TENCON 2018-2018 IEEE Region 10 Conference*, pages 0657–0662. IEEE.

Varsha D Badal and Colin A Depp. 2022. Natural language processing of medical records: new understanding of suicide ideation by dementia subtypes: Commentary on "suicidal ideation in dementia: associations with neuropsychiatric symptoms and subtype diagnosis" by naismith et al. *International Psychogeriatrics*, 34(4):319–321.

Loïc Barrault, Yu-An Chung, Mariano Cora Meglioli, David Dale, Ning Dong, Paul-Ambroise Duquenne, Hady Elsahar, Hongyu Gong, Kevin Heffernan, John

Hoffman, et al. 2023. Seamlessm4t-massively multilingual & multimodal machine translation. *arXiv preprint arXiv:2308.11596*.

Tom L Beauchamp et al. 2008. The belmont report. *The Oxford textbook of clinical research ethics*, pages 149–155.

Oumaima Benlaaraj, Ilyas El Jaafari, Ayoub Ellahyani, and Idriss Boutaayamou. 2022. Prediction of suicidal ideation in a new arabic annotated dataset. In *2022 9th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, pages 1–5. IEEE.

Marouane Birjali, Abderrahim Beni-Hssane, and Mohammed Erritali. 2017. Machine learning and semantic sentiment analysis based algorithms for suicide sentiment prediction in social networks. *Procedia Computer Science*, 113:65–72.

Jordan M Braciszewski. 2021. Digital technology for suicide prevention. *Advances in psychiatry and behavioral health*, 1(1):53–65.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Pete Burnap, Gualtiero Colombo, Rosie Amery, Andrei Hodorog, and Jonathan Scourfield. 2017. Multi-class machine classification of suicide-related communication on twitter. *Online social networks and media*, 2:32–44.

Cansu Canca. 2020. Operationalizing ai ethics principles. *Communications of the ACM*, 63(12):18–21.

Lei Cao, Huijun Zhang, Ling Feng, Zihan Wei, Xin Wang, Ningyun Li, and Xiaohao He. 2019. Latent suicide risk detection on microblog via suicide-oriented word embeddings and layered attention. *arXiv preprint arXiv:1910.12038*.

Sheila Castilho, Joss Moorkens, Federico Gaspari, Rico Sennrich, Andy Way, and Panayota Georgakopoulou. 2018. Evaluating mt for massive open online courses: A multifaceted comparison between pbsmt and nmt systems. *Machine translation*, 32(3):255–278.

Karen L Celedonia, Marcelo Corrales Compagnucci, Timo Minssen, and Michael Lowery Wilson. 2021. Legal, ethical, and wider implications of suicide risk detection systems in social media platforms. *Journal of Law and the Biosciences*, 8(1):lsab021.

Glen Coppersmith, Ryan Leary, Eric Whyne, and Tony Wood. 2015. Quantifying suicidal ideation via language usage on social media. In *Joint statistics meetings proceedings, statistical computing section, JSM*, volume 110.

Marta R Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, et al. 2022. No language left behind: Scaling human-centered machine translation. *arXiv preprint arXiv:2207.04672*.

Saloni Dattani, Lucas Rodés-Guirao, Hannah Ritchie, Max Roser, and Esteban Ortiz-Ospina. 2023. Suicides. *Our World in Data*. Https://ourworldindata.org/suicide.

Susan L Davis. 2023. Bloom: A multilingual open-source model for addressing ai text biases in natural language processing-a comprehensive review. *Advances in AI*, 1(1).

Lauren M Denneson, Holly B Williams, Mark S Kaplan, Bentson H McFarland, and Steven K Dobscha. 2016. Treatment of veterans with mental health symptoms in va primary care prior to suicide. *General Hospital Psychiatry*, 38:65–70.

Jingcheng Du, Yaoyun Zhang, Jianhong Luo, Yuxi Jia, Qiang Wei, Cui Tao, and Hua Xu. 2018. Extracting psychiatric stressors for suicide from social media using deep learning. *BMC medical informatics and decision making*, 18(2):77–87.

Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, et al. 2021. An ethical framework for a good ai society: Opportunities, risks, principles, and recommendations. *Ethics, governance, and policies in artificial intelligence*, pages 19–39.

Gottlob Frege. 1892. Über sinn und bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50.

Manas Gaur, Amanuel Alambo, Joy Prakash Sain, Ugur Kursuncu, Krishnaprasad Thirunarayan, Ramakanth Kavuluru, Amit Sheth, Randy Welton, and Jyotishman Pathak. 2019. Knowledge-aware assessment of severity of suicide risk for early intervention. In *The world wide web conference*, pages 514–525.

Norberto Nuno Gomes de Andrade, Dave Pawson, Dan Muriello, Lizzy Donahue, and Jennifer Guadagno. 2018. Ethics and artificial intelligence: suicide prevention on facebook. *Philosophy & Technology*, 31:669–684.

Naman Goyal, Cynthia Gao, Vishrav Chaudhary, Peng-Jen Chen, Guillaume Wenzek, Da Ju, Sanjana Krishnan, Marc'Aurelio Ranzato, Francisco Guzmán, and Angela Fan. 2021. The flores-101 evaluation benchmark for low-resource and multilingual machine translation.

Yvette Graham, Timothy Baldwin, Alistair Moffat, and Justin Zobel. 2017. Can machine translation systems be evaluated by the crowd alone. *Natural Language Engineering*, 23(1):3–30.

Mariam Hassib, Nancy Hossam, Jolie Sameh, and Marwan Torki. 2022. Aradepsu: Detecting depression and suicidal ideation in arabic tweets using transformers. In *Proceedings of the The Seventh Arabic Natural Language Processing Workshop (WANLP)*, pages 302–311.

Xiaolei Huang, Lei Zhang, David Chiu, Tianli Liu, Xin Li, and Tingshao Zhu. 2014. Detecting suicidal ideation in chinese microblogs with psychological lexicons. In *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*, pages 844–849. IEEE.

Andrew T Jebb, Vincent Ng, and Louis Tay. 2021. A review of key likert scale development advances: 1995–2019. *Frontiers in psychology*, 12:637547.

Smriti Jha, Gerry Chan, Rita Orji, et al. 2023. Identification of risk factors for suicide and insights for developing suicide prevention technologies: A systematic review and meta-analysis. *Human Behavior and Emerging Technologies*, 2023.

Shaoxiong Ji, Xue Li, Zi Huang, and Erik Cambria. 2022. Suicidal ideation and mental disorder detection with attentive relation networks. *Neural Computing and Applications*, 34(13):10309–10319.

Shaoxiong Ji, Shirui Pan, Xue Li, Erik Cambria, Guodong Long, and Zi Huang. 2020. Suicidal ideation detection: A review of machine learning methods and applications. *IEEE Transactions on Computational Social Systems*, 8(1):214–226.

Junglyun Kim, DongHyeon Gwak, Seonhee Kim, and Moonhee Gang. 2023. Identifying the suicidal ideation risk group among older adults in rural areas: Developing a predictive model using machine learning methods. *Journal of Advanced Nursing*, 79(2):641–651.

Olivia J Kirtley, Kasper van Mens, Mark Hoogendoorn, Navneet Kapur, and Derek de Beurs. 2022. Translating promise into practice: a review of machine learning in suicide research and prevention. *The Lancet Psychiatry*, 9(3):243–252.

Michael Kwet. 2019. Digital colonialism: Us empire and the new imperialism in the global south. *Race & Class*, 60(4):3–26.

Mark E Larsen, Nicholas Cummins, Tjeerd W Boonstra, Bridianne O'Dea, Joe Tighe, Jennifer Nicholas, Fiona Shand, Julien Epps, and Helen Christensen. 2015. The use of technology in suicide prevention. In *2015 37th annual international conference of the IEEE engineering in Medicine and biology society (EMBC)*, pages 7316–7319. IEEE.

Daeun Lee, Migyeong Kang, Minji Kim, and Jinyoung Han. 2022. Detecting suicidality with a contextual graph neural network. In *Proceedings of the eighth workshop on computational linguistics and clinical psychology*, pages 116–125.

Daeun Lee, Soyoung Park, Jiwon Kang, Daejin Choi, and Jinyoung Han. 2020. Cross-lingual suicidal-oriented word embedding toward suicide prevention. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2208–2217.

William Lewis, Robert Munro, and Stephan Vogel. 2011. Crisis mt: Developing a cookbook for mt in crisis situations. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 501–511.

Todd D Little, Kathleen M Roche, Sy-Miin Chow, Anna P Schenck, and Leslie-Ann Byam. 2016. National institutes of health pathways to prevention workshop: Advancing research to prevent youth suicide. *Annals of internal medicine*, 165(11):795–799.

Meizhen Lv, Ang Li, Tianli Liu, and Tingshao Zhu. 2015. Creating a chinese suicide dictionary for identifying suicide risk on social media. *PeerJ*, 3:e1455.

Laura Martinengo, Louise Van Galen, Elaine Lum, Martin Kowalski, Mythily Subramaniam, and Josip Car. 2019. Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. *BMC medicine*, 17(1):1–12.

Thomas H McCoy, Victor M Castro, Ashlee M Roberson, Leslie A Snapper, and Roy H Perlis. 2016. Improving prediction of suicide and accidental death after discharge from general hospitals with natural language processing. *JAMA psychiatry*, 73(10):1064–1071.

Brian L Mishara and David N Weisstub. 2016. The legal status of suicide: A global review. *International journal of law and psychiatry*, 44:54–74.

Rohan Mishra, Pradyumn Prakhar Sinha, Ramit Sawhney, Debanjan Mahata, Puneet Mathur, and Rajiv Ratn Shah. 2019. Snap-batnet: Cascading author profiling and social network graphs for suicide ideation detection on social media. In *Proceedings of the 2019 conference of the North American Chapter of the association for computational linguistics: student research workshop*, pages 147–156.

Destinee Morrow, Rafael Zamora-Resendiz, Jean C Beckham, Nathan A Kimbrel, David W Oslin, Suzanne Tamang, Silvia Crivelli, Million Veteran Program Suicide Exemplar Work Group, et al. 2022. A case for developing domain-specific vocabularies for extracting suicide factors from healthcare notes. *Journal of psychiatric research*, 151:328–338.

Yasmin Moslem, Rejwanul Haque, and Andy Way. 2023. Adaptive machine translation with large language models. *arXiv preprint arXiv:2301.13294*.

Yasmin Moslem, Andy Way, Rejwanul Haque, and John D Kelleher. Domain-specific text generation for machine translation. *Volume 1: MT Research Track*.

T. Munzner.

Sharon O'Brien. 2017. Machine translation and cognition. *The handbook of translation and cognition*, pages 311–331.

Bridianne O'dea, Stephen Wan, Philip J Batterham, Alison L Calear, Cecile Paris, and Helen Christensen. 2015. Detecting suicidality on twitter. *Internet Interventions*, 2(2):183–188.

National Institutes of Health et al. 2016. Policy on good clinical practice training for nih awardees involved in nih-funded clinical trials. *NOT-OD-16-1482017*.

Elena Okhapkina, Valentin Okhapkin, and Oleg Kazarin. 2017. Adaptation of information retrieval methods for identifying of destructive informational influence in social networks. In *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, pages 87–92. IEEE.

World Health Organization et al. 2021. Suicide worldwide in 2019: global health estimates.

Martin Orr, Kirsten Van Kessel, and David Parry. 2022. The ethical role of computational linguistics in digital psychological formulation and suicide prevention. In *Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology*.

John E. Ortega and Kenneth W. Church. 2023. A research-based guide for the creation and deployment of a low-resource machine translation system. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2023)*, pages 809–819.

Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. 2019. fairseq: A fast, extensible toolkit for sequence modeling. *arXiv preprint arXiv:1904.01038*.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.

Hilary Putnam. 1981. *Reason, truth and history*, volume 3. Cambridge University Press.

Diana Ramírez-Cifuentes, Ana Freire, Ricardo Baeza-Yates, Joaquim Puntí, Pilar Medina-Bravo, Diego Alejandro Velazquez, Josep Maria Gonfaus, and Jordi Gonzàlez. 2020. Detection of suicidal ideation on social media: multimodal, relational, and behavioral analysis. *Journal of medical internet research*, 22(7):e17758.

Bertrand Russell. 1905. On denoting. *Mind*, 14(56):479–493.

Samer Muthana Sarsam, Hosam Al-Samarraie, Ahmed Ibrahim Alzahrani, Waleed Alnumay, and Andrew Paul Smith. 2021. A lexicon-based approach to detecting suicide-related messages on twitter. *Biomedical Signal Processing and Control*, 65:102355.

Ramit Sawhney, Harshit Joshi, Saumya Gandhi, and Rajiv Ratn Shah. 2021. Towards ordinal suicide ideation detection on social media. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pages 22–30.

Ramit Sawhney, Prachi Manchanda, Raj Singh, and Swati Aggarwal. 2018. A computational approach to feature extraction for identification of suicidal ideation in tweets. In *Proceedings of ACL 2018, Student Research Workshop*, pages 91–98.

Annika Marie Schoene, John Ortega, Silvio Amir, and Kenneth Church. 2023. An example of (too much) hyper-parameter tuning in suicide ideation detection. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 17, pages 1158–1162.

Pradyumna Prakhar Sinha, Rohan Mishra, Ramit Sawhney, Debanjan Mahata, Rajiv Ratn Shah, and Huan Liu. 2019. # suicidal-a multipronged approach to identify and explore suicidal ideation in twitter. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 941–950.

John Snowdon and Namkee G Choi. 2020. Undercounting of suicides: where suicide data lie hidden. *Global public health*, 15(12):1894–1901.

Jeff Speaks. 2010. Theories of meaning.

Timothy R Tangherlini. 2013. *Danish folktales, legends, and other stories*. University of Washington Press.

NLLB Team, Marta R Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, et al. 2022a. No language left behind: Scaling human-centered machine translation. *línea]. Disponible en: https://github.com/facebookresearch/fairseq/tree/nllb*.

NLLB Team, Marta R. Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, Anna Sun, Skyler Wang, Guillaume Wenzek, Al Youngblood, Bapi Akula, Loic Barrault, Gabriel Mejia Gonzalez, Prangthip Hansanti, John Hoffman, Semarley Jarrett, Kaushik Ram Sadagopan, Dirk Rowe, Shannon Spruit, Chau Tran, Pierre Andrews, Necip Fazil Ayan, Shruti Bhosale, Sergey Edunov, Angela Fan, Cynthia Gao, Vedanuj Goswami, Francisco Guzmán, Philipp Koehn, Alexandre Mourachko, Christophe Ropers, Safiyyah Saleem, Holger Schwenk, and Jeff Wang. 2022b. No language left behind: Scaling human-centered machine translation.

Kid Valeriano, Alexia Condori-Larico, and Josè Sulla-Torres. 2020. Detection of suicidal intent in spanish language social networks using machine learning. *International Journal of Advanced Computer Science and Applications*, 11(4).

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Ludwig Wittgenstein. 2010. *Philosophical investigations*. John Wiley & Sons.

Edward N Zalta, Uri Nodelman, Colin Allen, and John Perry. 1995. Stanford encyclopedia of philosophy.

A Appendix

In Table A we show all available languages for translation based on a list of languages from the Flores dataset (Goyal et al., 2021) used for the pre-trained language model.

Language	**Script**	**Res.**
Acehnese	Arabic	low
Acehnese	Latin	low
Mesopotamian Arabic	low	Arabic
Tunisian Arabic	Arabic	low
Afrikaans	Latin	high
South Levantine Arabic	Arabic	low
Akan	Latin	low
Amharic	Ge'ez	low
North Levantine Arabic	Arabic	low
Modern Standard Arabic	Arabic	high
Modern Standard Arabic	Latin	low
Najdi Arabic	Arabic	low
Moroccan Arabic	Arabic	low
Egyptian Arabic	Arabic	low
Assamese	Bengali	low
Asturian	Latin	low
Awadhi	Devanagari	low
Central Aymara	Latin	low
South Azerbaijani	Arabic	low
North Azerbaijani	Latin	low
Bashkir	Cyrillic	low
Bambara	Latin	low
Balinese	Latin	low
Belarusian	Cyrillic	low
Bemba	Latin	low
Bengali	Bengali	high
Bhojpuri	Devanagari	low
Banjar	Arabic	low
Banjar	Latin	low
Standard Tibetan	Tibetan	low
Bosnian	Latin	high
Buginese	Latin	low
Bulgarian	Cyrillic	high
Catalan	Latin	high
Cebuano	Latin	low
Czech	Latin	high
Chokew	Latin	low
Central Kurdish	Arabic	low
Crimean Tatar	Latin	low
Welsh	Latin	low
Danish	Latin	high
German	Latin	high
Southwestern Dinka	Latin	low
Dyula	Latin	low
Dzongkha	Tibetan	low
Greek	Greek	high
English	Latin	high
Esperanto	Latin	low
Estonian	Latin	high
Basque	Latin	high
Ewe	Latin	low
Faroese	Latin	low
Fijian	Latin	low
Finnish	Latin	high
Fon	Latin	low
French	Latin	high
Friulian	Latin	low
Nigerian Fulfulde	Latin	low
Scottish Gaelic	Latin	low
Irish	Latin	low
Galician	Latin	low
Guarani	Latin	low
Gujarati	Gujarati	low
Haitian Creole	Latin	low
Hausa	Latin	low
Hebrew	Hebrew	low
Hindi	Devanagari	high
Chhattisgarhi	Devanagari	low
Croatian	Latin	high
Hungarian	Latin	high
Armenian	Armenian	low
Igbo	Latin	low
Ilocano	Latin	low
Indonesian	Latin	high
Icelandic Latin	high	
Italian	Latin	high
Javanese	Latin	low
Japanese	Japanese	high
Kabyle	Latin	low

Jingpho	Latin	low	Northern Sotho	Latin	low
Kamba	Latin	low	Nuer	Latin	low
Kannada	Kannada	low	Nyanja	Latin	low
Kashmiri	Arabic	low	Occitan	Latin	low
Kashmiri	Devanagari	low	West Central Oromo	Latin	low
Georgian	Georgian	low	Odia	Oriya	low
Central Kanuri	Arabic	low	Pangasinan	Latin	low
Central Kanuri	Latin	low	Eastern Panjabi	Gurmukhi	low
Kazakh	Cyrillic	high	Papiamento	Latin	low
Kabiyè	Latin	low	Western Persian	Arabic	high
Kabuverdianu	Latin	low	Polish	Latin	high
Khmer	Khmer	low	Portuguese	Latin	high
Kikuyu	Latin	low	Dari	Arabic	low
Kinyarwanda	Latin	low	Southern Pashto	Arabic	low
Kyrgyz	Cyrillic	low	Ayacucho Quechua	Latin	low
Kimbundu	Latin	low	Romanian	Latin	high
Northern Kurdish	Latin	low	Rundi	Latin	low
Kikongo	Latin	low	Russian	Cyrillic	high
Korean	Hangul	high	Sango	Latin	low
Lao	Lao	low	Sanskrit	Devanagari	low
Ligurian	Latin	low	Santali	Ol Chiki	low
Limburgish	Latin	low	Sicilian	Latin	low
Lingala	Latin	low	Shan	Myanmar	low
Lithuanian	Latin	high	Sinhala	Sinhala	low
Lombard	Latin	low	Slovak	Latin	high
Latgalian	Latin	low	Slovenian	Latin	high
Luxembourgish	Latin	low	Samoan	Latin	low
Luba-Kasai	Latin	low	Shona	Latin	low
Ganda	Latin	low	Sindhi	Arabic	low
Luo	Latin	low	Somali	Latin	low
Mizo	Latin	low	Southern Sotho	Latin	high
Standard Latvian	Latin	high	Spanish	Latin	high
Magahi	Devanagari	low	Tosk Albanian	Latin	high
Maithili	Devanagari	low	Sardinian	Latin	low
Malayalam	Malayalam	low	Serbian	Cyrillic	low
Marathi	Devanagari	low	Swati	Latin	low
Minangkabau	Arabic	low	Sundanese	Latin	low
Minangkabau	Latin	low	Swedish	Latin	high
Macedonian	Cyrillic	high	Swahili	Latin	high
Plateau Malagasy	Latin	low	Silesian	Latin	low
Maltese	Latin	high	Tamil	Tamil	low
Meitei	Bengali	low	Tatar	Cyrillic	low
Halh Mongolian	Cyrillic	low	Telugu	Telugu	low
Mossi	Latin	low	Tajik	Cyrillic	low
Maori	Latin	low	Tagalog	Latin	high
Burmese	Myanmar	low	Thai	Thai	high
Dutch	Latin	high	Tigrinya	Ge‘ez	low
Norwegian Nynorsk	Latin	low	Tamasheq	Latin	low
Norwegian Bokmål	Latin	low	Tamasheq	Tifinagh	low
Nepali	Devanagari	low	Tok Pisin	Latin	low

Tswana	Latin	high
Tsonga	Latin	low
Turkmen	Latin	low
Tumbuka	Latin	low
Turkish	Latin	high
Twi	Latin	low
Central Atlas Tamazight	Tifinagh	low
Uyghur	Arabic	low
Ukrainian	Cyrillic	high
Umbundu	Latin	high
Urdu	Arabic	low
Northern Uzbek	Latin	high
Venetian	Latin	low
Vietnamese	Latin	high
Waray	Latin	low
Wolof	Latin	low
Xhosa	Latin	high
Eastern Yiddish	Hebrew	low
Yoruba	Latin	low
Yue Chinese	Han (Traditional)	low
Chinese	Han (Simplified)	high
Chinese	Han (Traditional)	high
Standard Malay	Latin	high
Zulu	Latin	high