



# Model Card

## Lelapa-X-NER (isiZulu)

<b>Model Details</b>	<b>1</b>
Intended use	2
Primary intended uses	2
Primary intended users	2
Out-of-scope use cases	2
Factors	2
Relevant factors	2
Evaluation factors	3
Metrics	3
Model performance measures	3
Decision thresholds	4
Approaches to Uncertainty and Variability	4
Evaluation data	4
Datasets	4
Motivation	4
Preprocessing	5
Training data	5
Quantitative analyses	5
Unitary results	5
Human evaluation	6
Intersectional result	7
Ethical considerations	7
Caveats and recommendations	7



## Model Details

*Basic information about the model:* Review section 4.1 of the [model cards paper](#).

Organization	Lelapa AI
Product	Vulavula
Model date	7 November 2023
Feature	ASR
Lang	isiZulu
Domain	News
Model Name	Lelapa-X-NER (isiZulu)
Model version	1.0.0
Model Type	Fine-Tuned Proprietary Model

**Information about training algorithms, parameters, fairness constraints or other applied approaches, and features:** Proprietary Fine-tuning of a Base Model on Text Data

**License:** Proprietary

**Contact:** [info@lelapa.ai](mailto:info@lelapa.ai)

## Intended use

*Use cases that were envisioned during development:* Review section 4.2 of the [model cards paper](#).

### Primary intended uses

Intended use is governed by the language and domain of the model. The model is intended to be used in the news domain for Named Entity Recognition in isiZulu. The model is not suitable for domains different from News Articles and should be used with caution.



---

## Primary intended users

The NER model can be used for :

- News Aggregation
- Research
- Document Analysis

## Out-of-scope use cases

All languages and domains outside of NER Classification for isiZulu.

## Factors

*Factors could include demographic or phenotypic groups, environmental conditions, technical attributes, or others listed in Section 4.3: Review section 4.3 of the [model cards paper](#).*

## Relevant factors

Groups:

- The annotators assigned and their level of understanding of the task is one of the relevant factors. There is no record of the demographic information about the annotators.

Environmental conditions, Instrumentation & Technical attributes:

- The NER tags used are limited to Personal name (PER), Location (LOC), Organization (ORG), date & time (DATE), and Other (O)

## Evaluation factors

- In our development setting (training and evaluation), we used the tags described above to determine model accuracy.



---

## Metrics

*The appropriate metrics to feature in a model card depend on the type of model that is being tested. For example, classification systems in which the primary output is a class label differ significantly from systems whose primary output is a score. In all cases, the reported metrics should be determined based on the model's structure and intended use: Review section 4.4 of the [model cards paper](#).*

### Model performance measures

The model is evaluated using F1 Score: the models' performances are measured by automatic metric. As an automatic metric, F1 Score is an harmonic mean of the Precision and Recall. Precision tells us how well the model avoids mistakenly labeling something as a named entity when it's not. Recall tells us how well the model captures all the actual named entities in the text. So a higher F1 score generally means the model is doing a good job of finding the right names and not mistakenly labeling other tokens [Read more](#).

**F1 score:** Testing on NER test set in isiZulu

### Decision thresholds

No decision thresholds have been specified

## Evaluation data

*All referenced datasets would ideally point to any set of documents that provide visibility into the source and composition of the dataset. Evaluation datasets should include datasets that are publicly available for third-party use. These could be existing datasets or new ones provided alongside the model card analyses to enable further benchmarking.*

Review section 4.5 of the [model cards paper](#).

### Datasets

- Publicly available isiZulu NER datasets in the News domain.



## Motivation

These datasets have been selected because they are open-source, high-quality, and cover the targeted languages . These help to capture interesting cultural and linguistic aspects that would be crucial in the development process for better performance.

## Training data

Review section 4.6 of the [model cards paper](#).

Refer to the datasheet provided

## Quantitative analyses

*Quantitative analyses should be disaggregated, that is, broken down by the chosen factors. Quantitative analyses should provide the results of evaluating the model according to the chosen metrics, providing confidence interval values when possible.*

Review section 4.7 of the [model cards paper](#).

## Unitary results

Entity	F1-Score
PER (Person)	0.96
ORG (Organization)	0.85
LOC (Location)	0.92
DATE	0.86



---

## Intersectional result

In progress

## Ethical considerations

*This section is intended to demonstrate the ethical considerations that went into model development, surfacing ethical challenges and solutions to stakeholders. The ethical analysis does not always lead to precise solutions, but the process of ethical contemplation is worthwhile to inform on responsible practices and next steps in future work: Review section 4.8 of the [model cards paper](#).*

All call center data is synthetic and so the model does not contain any personal information. More details in the datasheet.

## Caveats and recommendations

*This section should list additional concerns that were not covered in the previous sections.*

Review section 4.9 of the [model cards paper](#).

Additional caveats are outlined extensively in our Terms and Conditions.