

# Dictionary 4.0: Alternative Presentations for Indonesian Multilingual Dictionaries

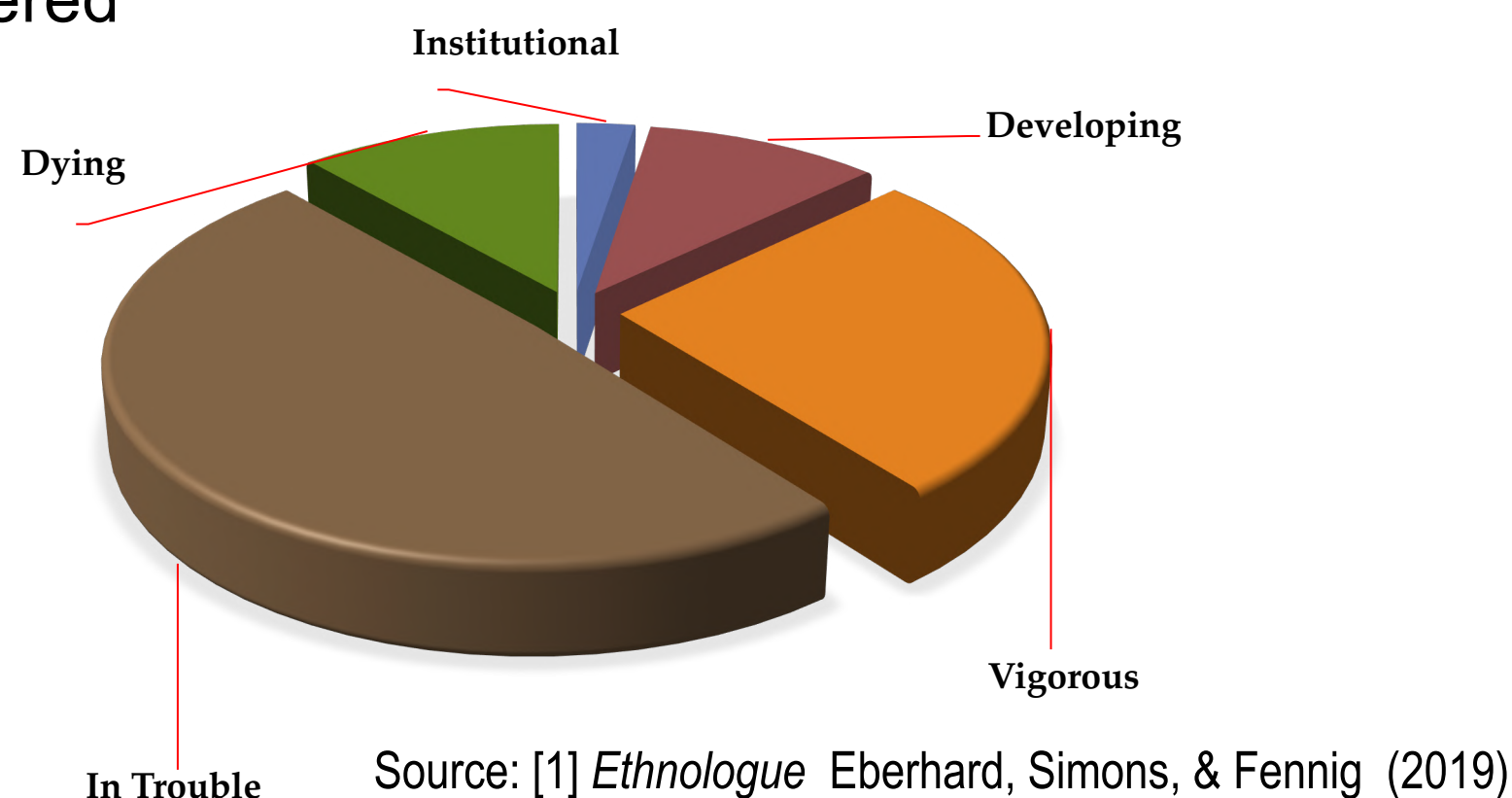
## Overview

- Building a multilingual dictionary for **719 languages** in Indonesia is a **challenging** task.
- Leipzig-Jakarta list** database for all indigenous languages in Indonesia
- Exploration** of the existing translations and **contribution** by **adding** translations to a new language or **editing** existing translations through **crowdsourcing**.

## Introduction

### Background

- More than half of Indonesian indigenous languages are endangered



- For starter, we covered 11 languages: Indonesian, Javanese, Sundanese, Madurese, Bima, Ternate, Tidore, Palembang Malay, Mandailing Batak, Malay, and Minangkabau.

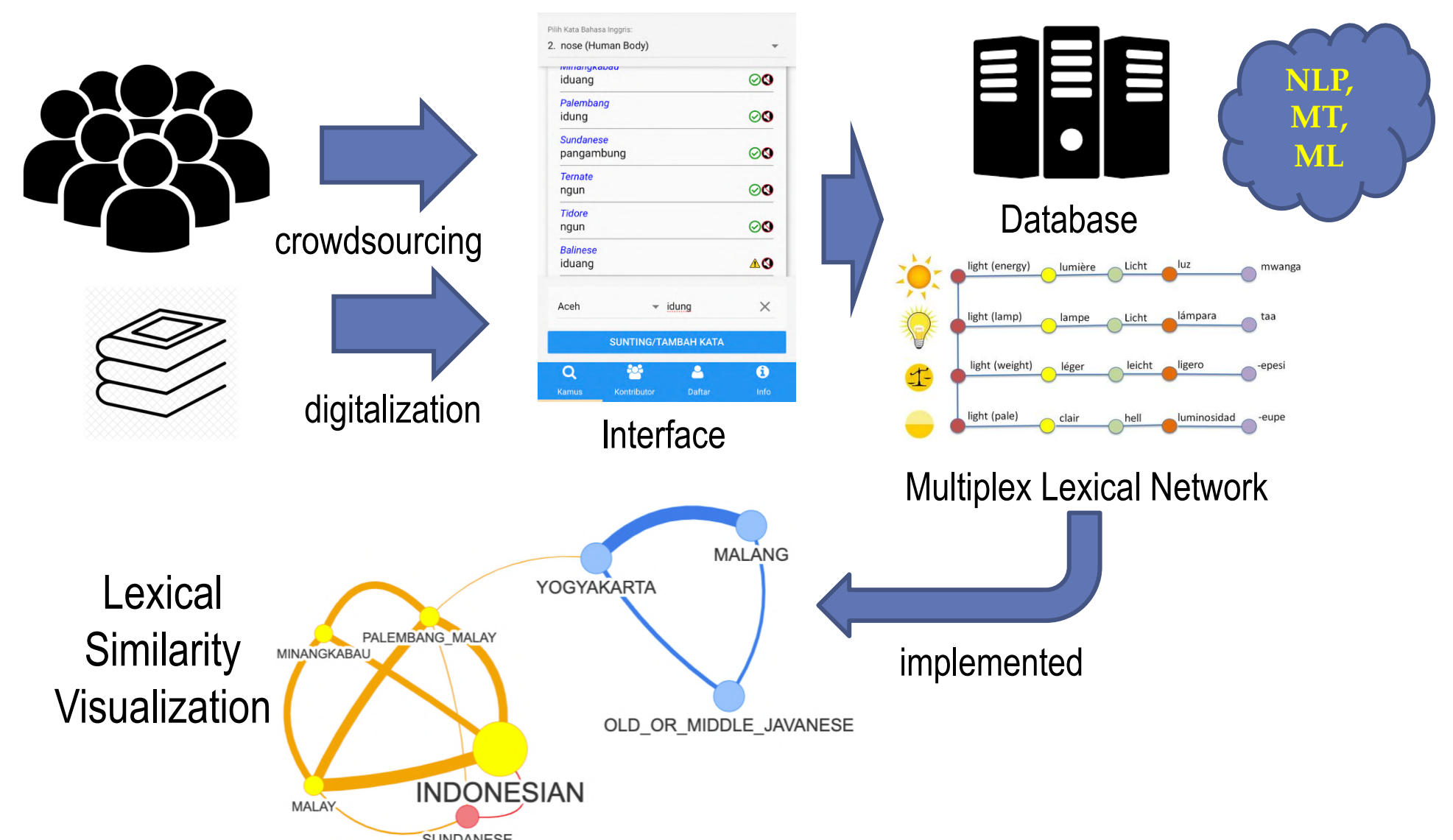
### Objective

- Building Leipzig-Jakarta list database for all indigenous languages in Indonesia
  - Language documentation & description
  - Implementation: generate lexical similarity or lexical distance matrix between languages by comparing the word list.

## Research Methodology

- Language Documentation + Language Description
- Language Documentation:**
  - Goal: producing grammar, dictionaries, and collections of texts,
  - the target users are generally linguists / researchers,
- Language Description:**
  - discourse-centered,
  - Goal: arrange a direct representation of a language / dialect with as many types of discourse as possible
- Swadesh List [2] vs Leipzig-Jakarta list [3]**
- Crowdsourcing + Gamification**
- The data will be used to generate **visualization** of Indonesian Indigenous Languages Lexical Similarity with Knowledge Graph

## Proposed Framework

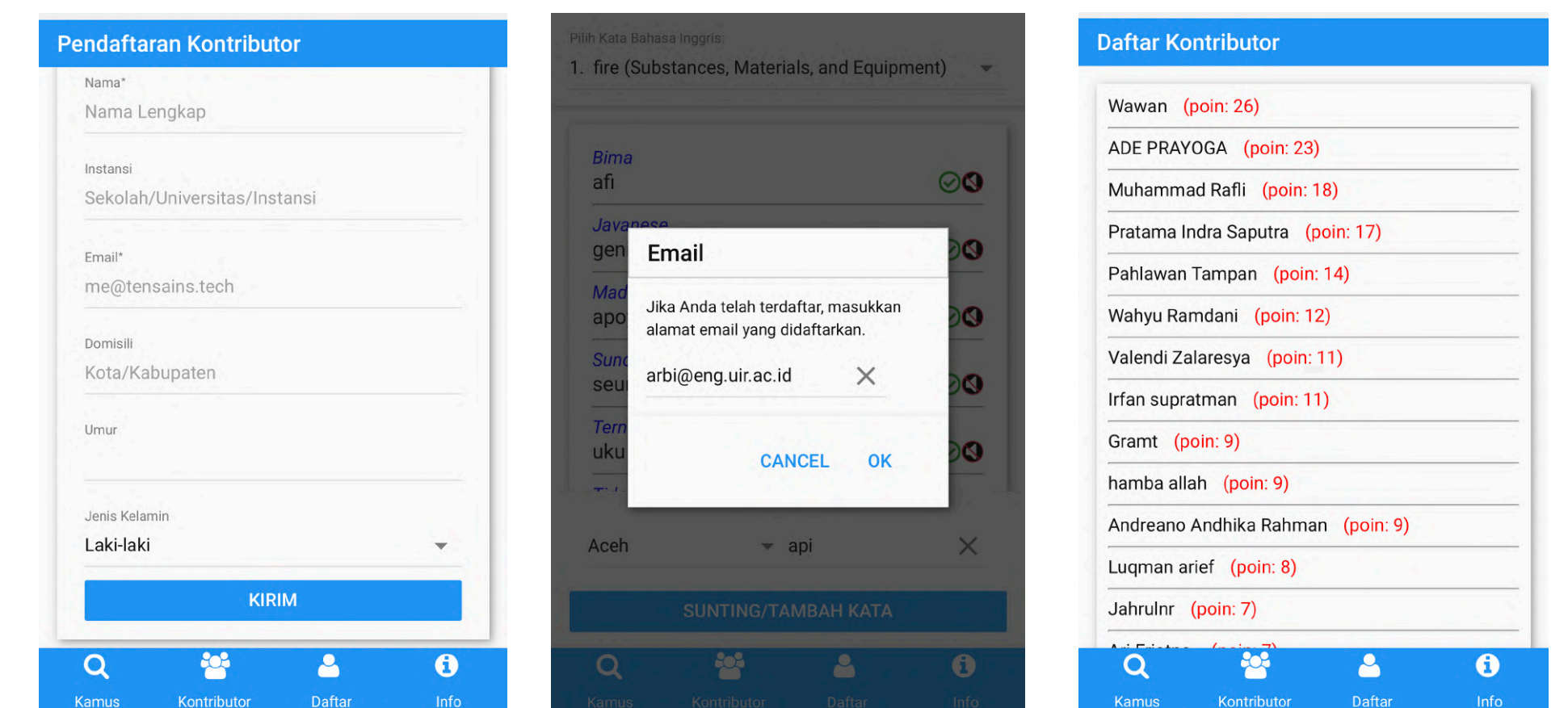


## Result

### Exploration

&

### Contribution




## Conclusion

Dictionary 4.0 is an easy to use apps for supporting Indonesian indigenous languages documentation and description with a good user acceptance (3.48/4).


## References

- [1] Eberhard, David M., Gary F. Simons, dan Charles D. Fennig (eds.). 2019. Ethnologue: Languages of the World. Twenty-second edition. Dallas, Texas: SIL International. Versi daring: <http://www.ethnologue.com>.
- [2] Swadesh, M., 1955. Towards greater accuracy in lexicostatistic dating. International journal of American linguistics, 21(2), pp.121-137.
- [3] Sakel, J., & Everett, D. L. 2012. Linguistic fieldwork: A student guide. Cambridge University Press.


## Researchers




\*Arbi Haza Nasution



\*\*Totok Suhardijanto



\* Informatics Engineering Department  
Faculty of Engineering, Universitas Islam Riau  
arbi@eng.uir.ac.id



\*\* Linguistics Department  
Faculty of Cultural Sciences, Universitas Indonesia  
totok.suhardijanto@ui.ac.id

**Acknowledgement**  
This research is supported by Universitas Islam Riau