



# TRANSLATOR GATOR: CROWDSOURCING TRANSLATION OF DEVELOPMENT KEYWORDS IN INDONESIA

PROGRAMME AREA: THE SUSTAINABLE DEVELOPMENT GOALS (SDGs)

TECHNOLOGY READINESS LEVEL: READY FOR SCALE OUT (LEVEL 6)

## SUMMARY

Pulse Lab Jakarta developed Translator Gator, a people-powered language game that creates dictionaries for recognising sustainable development-related conversations in Indonesia. The game builds taxonomies, i.e. sets of relevant keywords, by incentivising players to translate words from English into different Indonesian languages, including Bahasa Indonesia, Jawa, Sunda, Minang, Bugis and Melayu. During the pilot project, Translator Gator received over 109,000 contributions from 371 regular users across Indonesia. An initial analysis of the Bahasa Indonesia taxonomy showed crowdsourced translations were over 97 per cent accurate. Pulse Lab Jakarta gained significant insights into user preferences and incentives, which will be incorporated into a second phase of Translator Gator, that will be launched by the end of 2017.

## BACKGROUND

An essential element of public policy research is monitoring and analysing citizens’ feedback, both active and passive. Social media possesses a wealth of insights that can be mined to understand what people think and how they feel about matters affecting their lives. An important step in analysing social media data is creating taxonomies, i.e. sets of keywords, that can be used to extract insights relevant to topics of interest such as education, healthcare or early warning response.

Indonesia’s rich linguistic and cultural diversity, including some 300 different languages by some estimates, can pose difficulties for policymakers looking to gauge public opinion. Without a taxonomy of keywords to monitor, researchers are limited in their capacity to capture public discourse. On social media, such variations, including various dialects and jargon, make building a list of keywords more challenging as words, context and meaning change from region to region.

To address this challenge, Pulse Lab Jakarta developed Translator Gator, a web-based language game that crowdsources translations of words from English into any of Indonesia’s six most spoken languages. Translator Gator was inspired by the need to socialise the 17 Sustainable Development Goals (SDGs) and monitor progress against the varied indicators. The game aims to raise awareness of the SDGs among Indonesian youth and to develop a taxonomy of keywords to inform linguistic and social research.

## CROWDSOURCING THE TRANSLATION OF DEVELOPMENT KEYWORDS

Piloted from January to May 2016, Translator Gator asked players to translate SDG/development-related keywords and phrases from English into six Indonesian languages: Bahasa Indonesia, Jawa, Sunda, Bugis, Melayu, and Minang.

There are four components of the game, which are assigned randomly to players:

1. Translate SDG-related words/phrases from English to everyday (dictionary and non-dictionary) words/phrases in commonly used Indonesian languages;
2. Evaluate the translations submitted by others to validate the meaning;
3. Suggest synonyms, i.e. alternative word/phrases if necessary, and;
4. Classify words/phrases into selected categories outlined in the game.

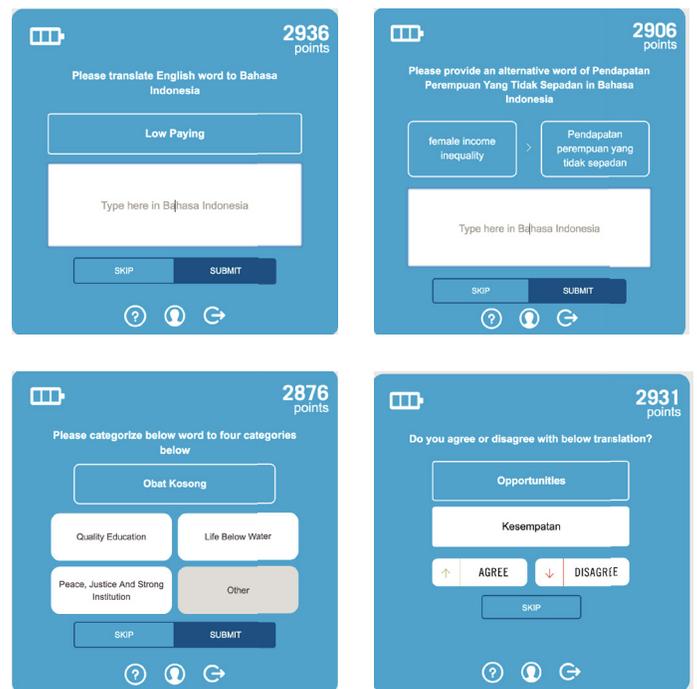


Image 1. The four game components of Translator Gator: translation, synonym, categorisation and evaluation.

The game employs two mechanisms to maintain player motivation and translation quality. First, players earn points for completing tasks and can exchange them for phone credit through a referral system. Second, a peer review mechanism encourages players to maintain a high quality of translations and skip translating words when unsure. For example, if a player's translations are disputed by others, then their "battery level" decreases. When the battery is empty, players need to wait an hour until it refills to be able to play again.

## INSIGHTS AND OUTCOMES

Using the Translator Gator platform, Pulse Lab Jakarta crowdsourced the translation of 1,609 words or phrases related to sustainable development into the six designated languages. During the pilot phase, the game received over 109,000 contributions from 371 regular players, creating a wealth of linguistic data. Among the 371 active users who played more than twice, 55% were female and 43% were male (while 2% did not reveal their gender). Most users were between the ages of 20 and 24.



Image 2. Translator Gator 2016 pilot phase statistics.

Player surveys revealed that the game's phone credit incentive was a major driver behind the project's successful data collection. Player contributions came in surges, which correlated with the redeem periods when players could exchange game points for phone credit.

The community engaged extensively in translations into Indonesian and Bahasa Jawa, but less so for the other local languages. An initial analysis of the Bahasa Indonesia taxonomy indicated that the crowdsourced translations were over 97 per cent accurate, confirming the 'wisdom of the crowd' when it comes to developing taxonomies and validating this approach for the purpose of the project.

During the pilot phase, Pulse Lab Jakarta gained insights into player preferences and patterns of use. For example, the analysis showed that players completed the second task—evaluating translations submitted by others—in the shortest amount of time, which indicated that the task is relatively easy. However, players indicated that evaluation was not the most preferred task because it was not challenging enough. In terms of timing preferences, participants were found to play more in the middle of the week and during break times, e.g. lunch time or after working hours.

## CONCLUSIONS

The project demonstrated the potential for using a crowdsourcing platform to build a taxonomy that supports localised public policy and linguistics research. The game produced high quality translations, which validated the crowdsourcing method and the quality control mechanisms and peer review features.

Pulse Lab Jakarta plans to launch a second phase of Translator Gator to inform disaster management efforts. The second phase will incorporate lessons learned from the initial pilot by enhancing some functionalities and improving player engagement through broader communication channels.

## IMPLICATIONS AND RECOMMENDATIONS

- This project demonstrates the potential of tapping into the "wisdom of the crowd" to create taxonomies which inform research on issues of relevance to sustainable development in Indonesia.
- The work conducted by Pulse Lab Jakarta would benefit from expanded project partnerships to leverage the knowledge, domain expertise and networks of local organisations and institutions.
- To improve socialisation efforts and expand the taxonomy completion rates in niche languages, Pulse Lab Jakarta will consider linking with relevant organisations and networks to target specific linguistic communities and user profiles.

## HOW TO CITE THIS DOCUMENT:

UN Global Pulse, 'Translator Gator: Crowdsourcing Translation of Development Keywords in Indonesia', Tool Series no. 4, 2017.