



Our Method--and How You Can Use It, Too

- With the support of the World Bank ECASTAT Trust Fund, we created an **open source method** aligned with the SDMX-SDG WG.
- With Rosstat and UNECE, our pilot study found reduced translation time by **40 percent**. The pilot was expanded to **all tier 1** metadata.
- StatCan and ECLAC/IADB were able to use our method to produce French and Spanish translations, with **minimal training** needed.

This is a description of our project--and **how you can use it, too**.



Translations Are Available to Anyone

- Translations are posted on our **website** as they are completed.

Концепции SDMX метаданных

- [0.a. Цель](#)
- [0.b. Задача](#)
- [0.c. Показатель](#)
- [0.g. Международные организации, ответственные за глобальный мониторинг](#)
- [2.a. Определение и понятия](#)
- [4.a. Обоснование](#)
- [4.b. Комментарии и ограничения](#)
- [4.c. Метод расчета](#)
- [4.f. Обработка отсутствующих](#)

Показатель: 1.1.1a

[Скачать PDF](#) [Revision history](#)

Цель 1: Повсеместная ликвидация нищеты во всех ее формах

Задача 1.1: К 2030 году ликвидировать крайнюю нищету для всех людей во всем мире (в настоящее время крайняя нищета определяется как проживание на сумму менее чем 1,25 долл. США в день)

Показатель 1.1.1: Доля населения, живущего за международной чертой бедности, в разбивке по полу,



Translations are Available in Several Languages

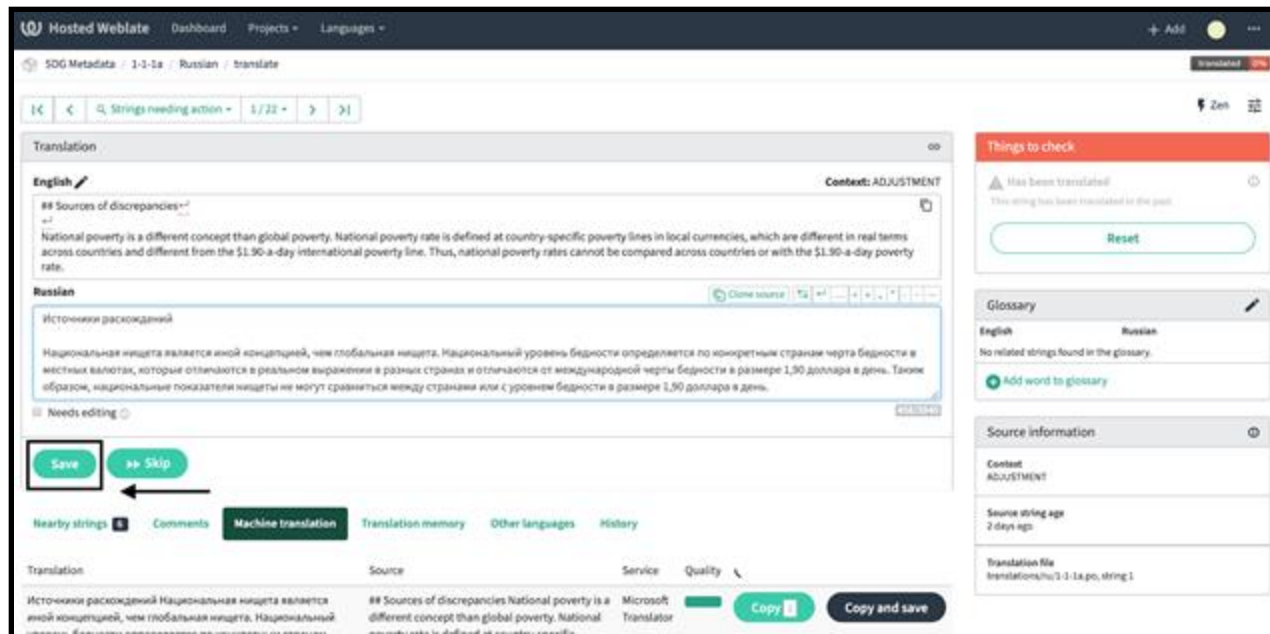
- Multiple translators can **work at their own pace**.





Translation Tools are Easy to Use and Open-Source

- The human translator uses Weblate to **translate** metadata.
- S/he can **select** from many translations and edit.





Revisions Can Be Made Easily

- Text to be changed is flagged for translators.

The screenshot shows a translation management interface for 'SDG Metadata / Site / Russian / Translate'. The interface includes navigation controls (4/4), a filter for 'Translated strings' with the state 'state:>=translated', and a 'Position and priority' dropdown. The main content is divided into three sections: 'Source change', 'English', and 'Russian'. The 'Source change' section shows a Latin text with a red highlight on 'Quisque interdum massa id leo varius posuere' and a green highlight on 'Proin cursus sapien quis molestie volutpat'. The 'English' section shows the translated text with a 'Key TESTING' label. The 'Russian' section shows the translated text with a 'Clone source' button and a 'NBS' label. At the bottom, the 'Review state' is set to 'Needs editing' (radio button selected), with options for 'Waiting for review' and 'Approved'. A counter '453/4340 - 434' is visible in the bottom right corner.



Other Translators Can Use Our Process

- Use our guides to translate into hundreds of other languages.

SDG Indicator Metadata Translation Project	
Step by Step Instructions	
Step 3: Instructions for Human Translator	
Updated December, 2020	
Contents	
<u>INSTRUCTIONS AT A GLANCE</u>	
<u>I. Basic Overview</u>	3
<u>II. Detailed Instructions</u>	4
<u>A. How to use Weblate</u>	4
<u>B. Formatting codes</u>	13



Translators' Experiences and Suggestions: Summary

- Translation time reduced by 30-60%
- Training materials easy to use and sufficient
- Weblate layout changes would improve convenience
- ★ Improve original English text for transparency and grammar

Next slides provide specific comments and details.



Translators' Experiences: Translation Time

From the pilot test (~20 metadata files), ROSSTAT estimated a 40% reduction in translation time using Weblate compared to traditional methods.

When the original text is strong, the translation is completely automated, with minor editing here and there. I would say in these cases, the time needed is cut by half to two third needed.(STATCAN)

The greatest time savings happen once the system learns and incorporates translations already made in other indicators...the last batch of indicators..required 3 times less time. (IDB-ECLAC)



Translators' Experiences: Training

I used the training material at the beginning to have a quick start on how to do things. After a quick read through, I was able to find my way in the system as I was sufficiently sure of what I needed to do. (STATCAN)

I used the training material and it was very helpful. My training was adequate and I have received clarification every time I have needed it. I would not recommend any changes, the translator should be capable enough to get the job done within this framework. (IDB-ECLAC)



Translators' Experiences: Tables, Images, and Formulas

It is desirable to make work with tables and formulae more transparent and understandable (without nightmares).(ROSSTAT)

Most of the tables were properly translated. The only issue here would be when you have a long dynamic hyperlink in the text. Their length skew the look of the proposed translations, making it extremely difficult to choose and edit the text. *[WB note: These are image codes where tables are so complex that they must be taken as images.]*



Translators' Priorities: Weblate Layout

It is more convenient for work when the original and translation texts are located side by side, not one below the other. (ROSSTAT)

Would it be possible to limit the number of translation robots and the size of the text [strings]? Some section of text can become quite large. Having a very long section cut in multiple sub-sections (1 or 2 small paragraphs) would make the translation process much easier and faster (STATCAN)

Baidu and Apertium Apy translation source are useless. Microsoft and Google translators are the most reliable sources..DeepL offers very good translations, but it does not always transcribe HTML codes correctly and sometimes has significant errors. (IDB-ECLAC)



Translators' Priorities: Improve English

If it is possible to address to authors of the texts, to write more simply and clearly. For example, in one indicator, the Dissemination is indicated as "From NA to NA". What does it mean? (ROSSTAT)

I would strongly suggest doing a light editing of the original text in English to ensure better continuity and standardization of the content. Beside improving the readability, this will sharply improve the quality of the translated documents and allow the system to learn from better material. (STATCAN)

Acronyms are not always detected / well translated. (IDB-ECLAC)



Next Steps: Tier 2 Indicators and More Languages

- Automate a **routine harvest** of metadata from UNSD's **SDG Lab**.
- Continuously **improve experience** with Weblate.
- Identify sponsors and collaborators for **additional languages**
- **Interested?**
 - Attend our World Bank Webinar **on April 27th**.
 - See our **website** at <https://worldbank.github.io/sdg-metadata/>



SDG METADATA TRANSLATION PROJECT

Using **Open-Source Solutions**
to Accelerate **SDG Metadata Translation**
and Build **National Statistical Capacity**



WORLD BANK GROUP

**SDG Metadata
Translation Team**

Contact us: sdgmetadatatranslation@worldbank.org

April 1, 2021