

Использование программного обеспечения с открытым исходным кодом для ускорения перевода метаданных ЦУР и наращивания национального статистического потенциала



Группа по переводу метаданных ЦУР

Свяжитесь с нами: sdgmetadatatranslation@worldbank.org

1 апреля 2021 года



#### Наш метод: как вы можете его использовать

- При поддержке Трастового фонда Всемирного банка (ECASTAT) мы создали метод с открытым исходным кодом, согласованный с РГ по SDMX-ЦУР.
- С помощью Росстата и ЕЭК ООН наше пилотное исследование показало сокращение времени перевода на **40 процентов**. Пилотный проект охватил все метаданные уровня **1**.
- Статистическая служба Канады и ЭКЛАК/МАБР смогли использовать наш метод для осуществления перевода на французский и испанский языки с минимальной необходимой подготовкой.

Перед вами описание нашего проекта и то, как вы можете его использовать.







• Переводы публикуются на нашем сайте по мере их завершения.

#### метаданных

0.а. Цель

<u>0.b. Задача</u>

0.с. Показатель

0.g. Международные организации, ответственные за глобальный мониторинг

2.а. Определение и понятия

4.а. Обоснование

4.b. Комментарии и ограничения

4.с. Метод расчета

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)

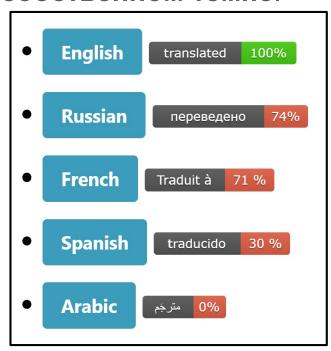
Показатель 1.1.1: Доля населения, живущего за





### Переводы доступны на нескольких языках

• Несколько переводчиков могут работать в своем собственном темпе.

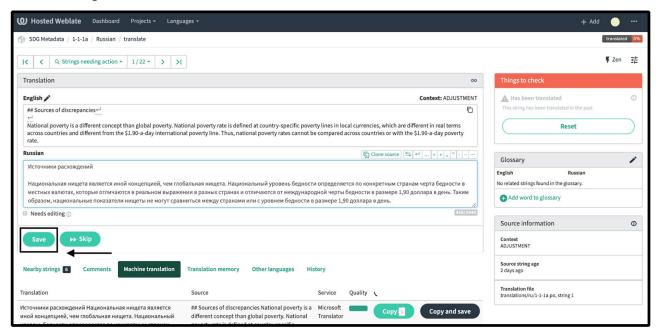




### Инструменты для перевода просты в использовании имеют открытый исходный код



- Переводчик использует веб-инструмент Weblate для перевода метаданных.
- Он/она может выбирать из множества вариантов переводов и редактировать их.

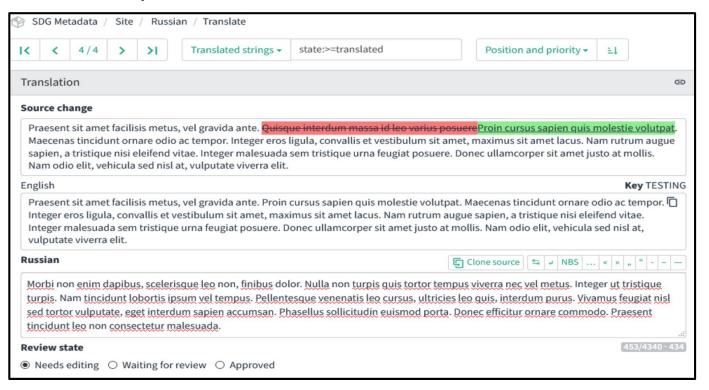






#### Можно легко вносить изменения

• Текст, который необходимо **изменить**, **выделяется** для удобства переводчиков.





### **Другие переводчики могут пользоваться нашими** ресурсами



• Используйте наши руководства для перевода на сотни других языков.

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





#### Опыт и предложения переводчиков: краткие выводы

- Время перевода сократилось на 30-60%
- Учебные материалы просты в использовании и подробно изложены
- Возможность менять оформление страницы на платформе
  Weblate была бы удобной
- ★ Улучшение оригинального английского текста для соблюдения понятности и грамматической точности

На следующих слайдах представлены конкретные комментарии и подробности.



### Опыт переводчиков: время, необходимое для выполнения перевода



По результатам пилотного теста (~20 файлов метаданных) РОССТАТ оценил сокращение времени перевода с помощью Weblate на 40% по сравнению с традиционными методами.

Если исходный текст корректный, то перевод выполняется полностью автоматически, и местами потребуется внести лишь незначительные изменения. Можно сказать, что в этих случаях необходимое время сокращается наполовину - на две трети (Статистическая служба Канады).

Наибольшая экономия времени происходит после того, как система изучает и применяет уже выполненные переводы других показателей. Для перевода последней партии показателей потребовалось в 3 раза меньше времени (МАБР-ЭКЛАК)





### Опыт переводчиков: обучение

Я использовал учебные материалы в самом начале, чтобы быстро понять, как что-то сделать. Бегло просмотрев материалы, я смог сориентироваться в системе, поскольку был достаточно уверен в том, что мне нужно было делать (Статистическая служба Канады)

Я использовал учебные материалы, и они были весьма полезны. Моя подготовка была достаточной, и я получал разъяснения каждый раз, когда они мне были необходимы. Я бы не рекомендовал никаких изменений, переводчик должен быть достаточно способным, чтобы выполнить работу в этой системе (МАБР-ЭКЛАК)



### Опыт переводчиков: таблицы, изображения и формулы

Желательно сделать работу с таблицами и формулами более доступной и понятной - без лишних мучений (РОССТАТ)

Большинство таблиц были переведены правильно. Единственная проблема здесь будет заключаться в том, что у вас есть длинная динамическая гиперссылка в тексте. Ее длина искажает внешний вид предлагаемых переводов, что крайне затрудняет выбор и редактирование текста. [Примечание WB: речь идет об изображениях кодов (image codes), на которых таблицы настолько сложны, что их нужно воспринимать как изображения.]



# Приоритеты переводчиков: оформление страницы на Weblate



Удобнее работать, когда тексты оригинала и перевода расположены рядом, а не один под другим (РОССТАТ)

Можно ли ограничить количество роботов-переводчиков и размер текста [строк]? Часть текста может быть довольно большой. В случае очень длинного куска текста, его можно было бы разделить на несколько подразделов (1 или 2 небольших абзаца), что значительно упростит и ускорит процесс перевода (Статистическая служба Канады)

Источники переводов Baidu и Apertium Apy бесполезны. Переводчики Microsoft и Google - самые надежные источники. Deepl предлагает очень хорошие переводы, но он не всегда правильно транскрибирует HTML-коды и иногда выдает существенные ошибки (МАБР-ЭКЛАК)



### Приоритеты переводчиков: улучшение английского языка

Если есть возможность обратиться к авторам текстов, то следует попросить их писать проще и понятнее. Например, в одном показателе слово "Dissemination" обозначено как «From NA to NA». Что это значит? (POCCTAT)

Я настоятельно рекомендую делать небольшое редактирование оригинального текста на английском языке, чтобы придать содержанию надлежащую целостность и упорядоченность. Помимо улучшения читабельности, это резко повысит качество переведенных документов и позволит системе учиться на более качественном материале (Статистическая служба Канады)

Аббревиатуры не всегда находятся/правильно переводены (МАБР-ЭКЛАК)



# Дальнейшие шаги: показатели уровня 2 и дополнительные языки

- Автоматизировать рутинный сбор метаданных из «Лаборатории» ЦУР Статистического отдела ООН.
- Постоянно повышать удобство использования Weblate.
- Находить спонсоров и сотрудников для обработки дополнительных языков
- Заинтересовались?
  - Посетите вебинар Всемирного банка 27 апреля.
  - Переходите на наш веб-сайт по ссылке: https://worldbank.github.io/sdg-metadata/



# ПРОЕКТ ПЕРЕВОДА МЕТАДАННЫХ ЦУР

Использование программного обеспечения с открытым исходным кодом для ускорения перевода метаданных ЦУР и наращивания национального статистического потенциала



Группа по переводу метаданных ЦУР

Свяжитесь с нами: sdgmetadatatranslation@worldbank.org

1 апреля 2021 года