



# ECASTAT TRUST FUND: SDG METADATA TRANSLATION PROJECT

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



**WORLD BANK GROUP**

April 27, 2021



# Background

- Since establishment in 2012, ECASTAT TF has been supporting the development of statistical organizations in the Europe and Central Asia region through improving methodological, technical, and human capacity.
- Capacity building is necessary for all countries to prepare the data and statistics needed to meet the scope and ambitions of the 2030 Agenda.
- When it comes to the Sustainable Development Goals (SDGs) monitoring and reporting, countries and teams are still requesting help and support for translating the metadata into local languages.



## Background (continued)

- Now more than ever, the transparency and accessibility of SDG data are essential for informing public and policymakers.
- SDG metadata (i.e., information on how to calculate and interpret SDG indicators) are not available languages other than English and Arabic. This has limited the understanding and implementation.
- By converting SDG metadata files into machine-readable format using open-source software, the team has developed an efficient and robust method for translation into different languages.

# Project requirements



- Preparing such translations is extremely time consuming and costly. There are **247** indicators (231 unique), **257** metadata files @ ~7 pages each of technical language.
- Version control is a challenge. The indicators are refined and revised over time, and it is difficult for users to know what changed and when.
- The method would have to be scaleable. Costs must be minimal, the tools must be easy to use, and a range of languages must be available.



# Strategic collaborations

- The ECASTAT support enabled the **data science work** to build and evaluate the translation protocol.
- UNECE funded translation of the indicator metadata files into **Russian** with expert appointed by ROSSTAT.
- UNSD's SDMX and IAEG-SDGs teams consulted regarding the **mapping of metadata concepts** in the preparation of machine-readable files.
- Additional collaboration and agreement were setup. Statistics Canada is preparing **French** and ECLAC with IDB has started the **Spanish** translation from the English machine-readable files.



# Three step approach

---

1. To make **machine-readable, SDMX compliant** SDG metadata files to speed and reduce costs of the translation process and aid version control.



# Three step approach

1. To make machine-readable, SDMX compliant SDG metadata files to speed and reduce costs of the translation process and aid version control.
2. To **translate** machine-readable metadata with computer-assisted translation software. A human translator with statistical expertise selects or edits the best translation for each metadata component.



# Three step approach

1. To make machine-readable, SDMX compliant SDG metadata files to speed and reduce costs of the translation process and aid version control.
2. To translate machine-readable metadata with computer-assisted translation software. A human translator with statistical expertise selects or edits the best translation for each metadata component.
3. To **post resulting translations** on the public website using open-source software. All files and documentation are available. This facilitates translation into additional languages by partners the reducing time and cost.





# Word file access

SDG tier 1 indicator files are accessed from UNSD Metadata Repository.

The screenshot shows the website [unstats.un.org/sdgs/metadata/](https://unstats.un.org/sdgs/metadata/). The page features a search sidebar on the left with the following elements:

- Search** (header)
- Enter Text (input field)
- Select Goal (dropdown menu)
- Select Target (dropdown menu)
- Filter (button)
- Clear (button)
- Back (button)

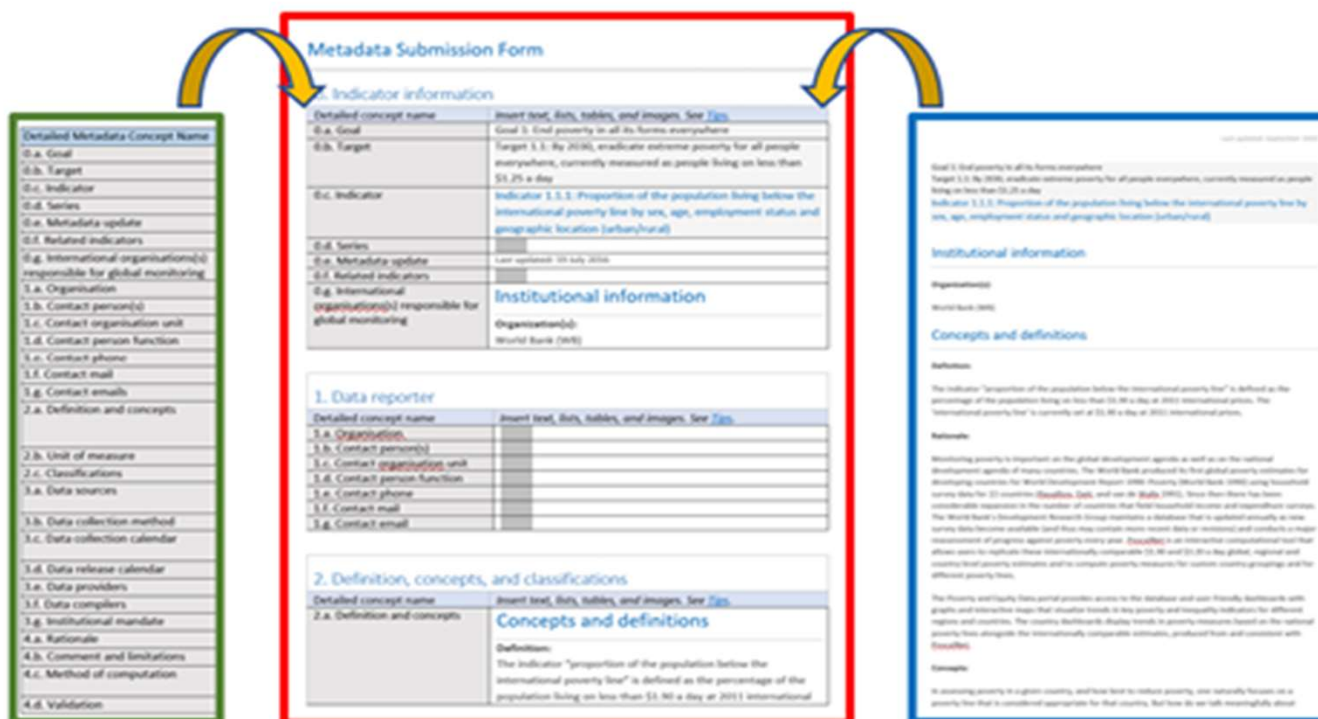
The main content area displays **Goal 1. End poverty in all its forms everywhere**. It lists three targets with their respective indicators and metadata links:

- Target 1.1:** By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day
  - Indicator 1.1.1:** Proportion of population below the international poverty line, by sex, age, employment status and geographical location (urban/rural)
    - See Metadata : (1)
    - See Metadata : (2)
- Target 1.2:** By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions
  - Indicator 1.2.1:** Proportion of population living below the national poverty line, by sex and age
    - See Metadata :
  - Indicator 1.2.2:** Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions
    - See Metadata :
- Target 1.3:** Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable



# Mapping of concepts

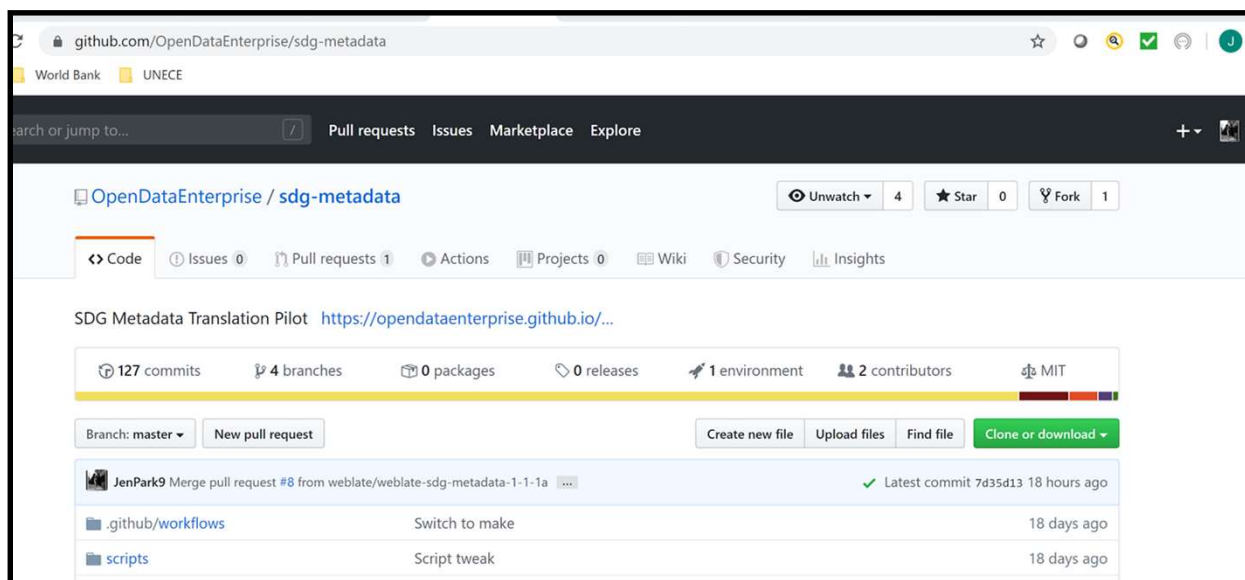
Team mapped the **SDMX metadata concepts** and the **IAEG-SDG metadata** into a machine-readable file, which allows automation.





# Transfer to Github

- Machine-readable metadata file are uploaded into **Github**, an open-source development platform.
- It is **free** for anyone to manage and share files.



# Easy, Open Source Translation Tools



- The human translator uses Weblate to **translate** metadata.
- S/he can **select from many translations and edit**.
- Add preferred translations to the **glossary**, improving efficiency.

The screenshot displays the Weblate web interface for translating a string. The main area shows the source text in English: "# Sources of discrepancies+  
National poverty is a different concept than global poverty. National poverty rate is defined at country-specific poverty lines in local currencies, which are different in real terms across countries and different from the \$1.90-a-day international poverty line. Thus, national poverty rates cannot be compared across countries or with the \$1.90-a-day poverty rate." Below this, the Russian translation is shown: "Источники расхождений  
Национальная нищета является иной концепцией, чем глобальная нищета. Национальный уровень бедности определяется по конкретным странам черта бедности в местных валютах, которые отличаются в реальном выражении в разных странах и отличаются от международной черты бедности в размере 1,90 доллара в день. Таким образом, национальные показатели нищеты не могут сравниться между странами или с уровнем бедности в размере 1,90 доллара в день." The interface includes a "Save" button (highlighted with a red box and an arrow pointing to the "Skip" button) and a "Skip" button. On the right side, there are panels for "Things to check" (indicating the string has been translated), "Glossary" (with an "Add word to glossary" button), "Source information" (showing context as ADJUSTMENT and source string age as 2 days ago), and a table of nearby strings with a "Copy and save" button.



# Revise Files Easily

- Text to be changed is flagged for translators.

SDG Metadata / Site / Russian / Translate

4 / 4 Translated strings state:>=translated Position and priority

Translation

**Source change**

Praesent sit amet facilisis metus, vel gravida ante. **Quisque interdum massa id leo varius posuere** **Proin cursus sapien quis molestie volutpat.** Maecenas tincidunt ornare odio ac tempor. Integer eros ligula, convallis et vestibulum sit amet, maximus sit amet lacus. Nam rutrum augue sapien, a tristique nisi eleifend vitae. Integer malesuada sem tristique urna feugiat posuere. Donec ullamcorper sit amet justo at mollis. Nam odio elit, vehicula sed nisl at, vulputate viverra elit.

English **Key TESTING**

Praesent sit amet facilisis metus, vel gravida ante. Proin cursus sapien quis molestie volutpat. Maecenas tincidunt ornare odio ac tempor. Integer eros ligula, convallis et vestibulum sit amet, maximus sit amet lacus. Nam rutrum augue sapien, a tristique nisi eleifend vitae. Integer malesuada sem tristique urna feugiat posuere. Donec ullamcorper sit amet justo at mollis. Nam odio elit, vehicula sed nisl at, vulputate viverra elit.

Russian [Clone source](#) NBS

Morbi non enim dapibus, scelerisque leo non, finibus dolor. Nulla non turpis quis tortor tempus viverra nec vel metus. Integer ut tristique turpis. Nam tincidunt lobortis ipsum vel tempus. Pellentesque venenatis leo cursus, ultricies leo quis, interdum purus. Vivamus feugiat nisl sed tortor vulputate, eget interdum sapien accumsan. Phasellus sollicitudin euismod porta. Donec efficitur ornare commodo. Praesent tincidunt leo non consectetur malesuada.

Review state 453/4340 · 434

Needs editing  Waiting for review  Approved



## Workflow Summary

- English metadata is entered into a **Word template**
- Word template is uploaded to the **Github repository**
- Github converts template to a **Gettext translation file**
- **Weblate pulls** the translation file from Github
- **Translators translate** the file into other languages
- **Weblate pushes** the translation files to Github
- Github updates the website to reflect the translations

# Results



- With Rosstat and UNECE, the pilot study found reduced translation time by **40 percent**.
- StatCan and ECLAC/IDB were able to use this method to produce French and Spanish translations, with **minimal training** needed.
- Process improvements are further **streamlining translation work** and **version control**.
- The project protocols, all machine-readable files in ENG, their translations, and training materials are available on the **public website** at <https://worldbank.github.io/sdg-metadata/>.



# Translations Are Available to Anyone



- Posted on the **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)

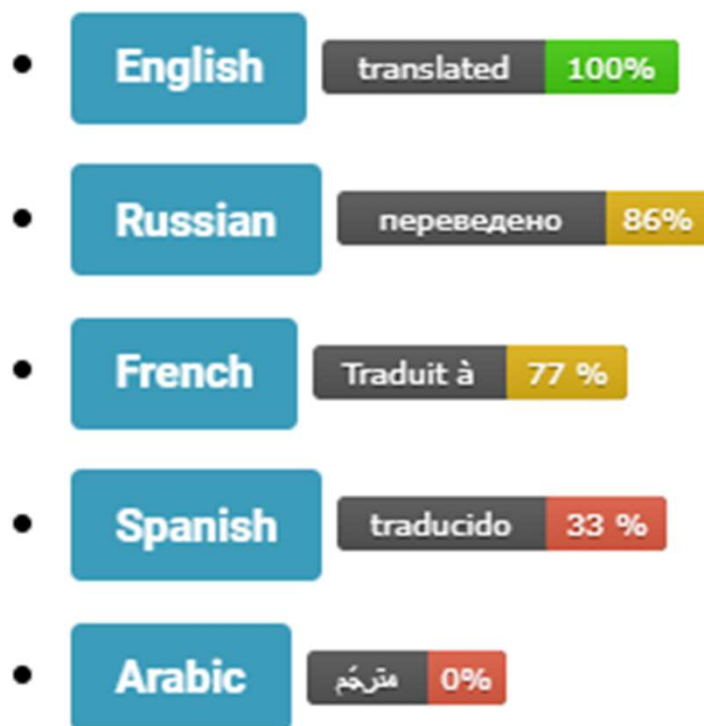
Показатель 1.1.1: Доля населения, живущего за международной чертой бедности, в разбивке по полу,



# Translations are Available in Several Languages



- Multiple translators can work at their own pace.



# Other Translators Can Use Project Process



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

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



# ECASTAT TRUST FUND: SDG METADATA TRANSLATION PROJECT

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



**WORLD BANK GROUP**

April 27, 2021