

# SDG Metadata Translation Project

Step by Step Instructions

## Step 1: Prepare Machine-Readable (and SDMX Compliant) Metadata Files

Updated December, 2020

---

### Contents

#### INSTRUCTIONS AT A GLANCE

##### I. CHOOSE WHICH METADATA FILE TO TRANSLATE

##### II. ACCESS UNSD SDG METADATA REPOSITORY

##### III. OPEN AUTHORING TEMPLATE

##### IV. CUT AND PASTE TEXT ACCORDING TO TEMPLATE MAPPING GUIDE

##### V. FORMAT PASTED TEXT USING GUIDANCE

##### VI. SAVE AS WORD FILE USING NAMING CONVENTION

##### VII. CHECK FORMATTING AS IT WOULD APPEAR IN MACHINE-READABLE FORMAT

---

### INSTRUCTIONS AT A GLANCE

**Purpose:** These are step-by-step instructions for ***Step 1, how to prepare machine-readable (and SDMX compliant) metadata files from Word documents***. Instructions for generating and managing machine-readable files in Github for translation into Weblate is described in Step 2. Step 3 describes how to use Weblate to translate machine-readable files and access the output.

**Scope:** For the SDG Metadata Translation Project, **Step 1** (described here) was completed for metadata pertaining to SDG indicators designated by the IAEG-SDG as tier 1, tier 1/tier2, or provisional tier 1 as of July 17, 2020. In total, this results in 135 files (some indicators are associated with more than one metadata file). Preparation of tier 2 indicator metadata files is pending approval.

**Users:** This work was completed by the World Bank and its consultants. **Human translators will not need to complete this step.**

**Needed materials:** The most recent IAEG-SDG approved version of SDG metadata files in either Word (preferred) or pdf format are needed. SDG indicators are revised and updated on a scheduled basis. See <https://unstats.un.org/sdgs/metadata/> for status of pending revisions.

**Questions?** Contact [sdgmetadatatranslation@worldbank.org](mailto:sdgmetadatatranslation@worldbank.org).

---

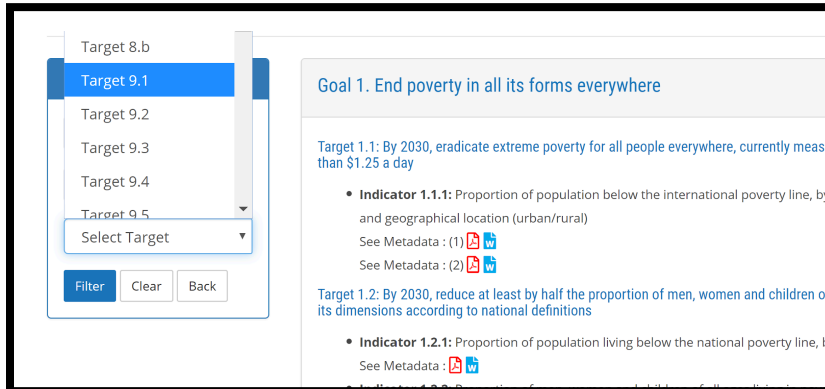
## I. CHOOSE WHICH METADATA FILE TO TRANSLATE

1. The SDG indicators and their metadata are periodically refined, revised, added, or deleted over time to reflect emerging science and improved methodologies. They are prepared by custodian agencies, proposed by the IAEG-SDGs and subsequently, by the UN Statistical Commission. **A list of the most recently approved indicators** can be found at <https://unstats.un.org/sdgs/indicators/indicators-list/>.
2. **Metadata for recently revised indicators may not yet be available on the UNSD website.** The March 2020 Comprehensive Review proposed a number of changes to SDG indicators and/or their metadata. See <https://unstats.un.org/sdgs/iaeg-sdgs/2020-comprev/UNSC-proposal/>. Check the date of last update on posted indicator metadata files (in the header of each metadata file) to make sure you are working with the expected, most recent version of the metadata file,
3. **We started with tier 1 indicators** (those with established globally agreed upon methodology and routinely collected data) . We used the tiering status as designated on 17 July 2020. We then worked with tier 1 indicator files as they were updated and posted on the UNSD website over summer/fall 2020. The tiering status of each indicator can be found at <https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/>).
4. **Metadata can also change even when the indicator has not.** Custodian agencies would submit these changes when they submit their SDG data to UNSD. Such changes need to be approved by the IAEG-SDG before they can be posted on the UNSD website. Therefore, we also suggest checking proposed revisions of metadata pending with UNSD at [statistics@un.org](mailto:statistics@un.org).
5. **SDG indicators and metadata will continue to change over time.** These revisions can be made easily to the [Word template](#) described in these instructions, converted into machine-readable form (see Step-by-Step Guide 2), and accordingly refreshed in Github and Weblate.

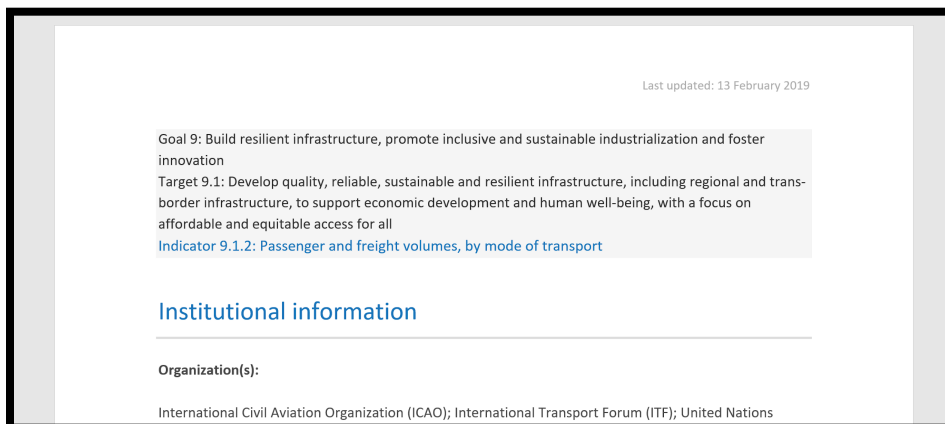
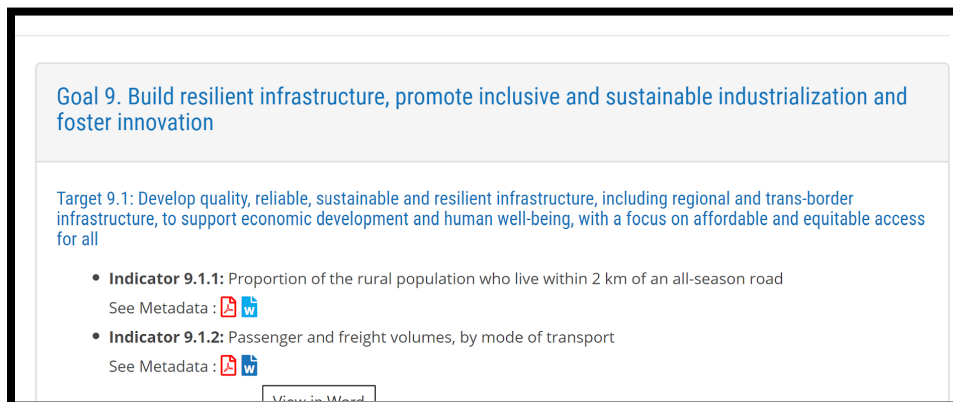
## II. ACCESS UNSD SDG METADATA REPOSITORY

6. Go to the UNSD SDG Metadata Repository at <https://unstats.un.org/sdgs/metadata/>.

7. Use the filter tool to select the indicator metadata file desired.



8. Click on the Word or pdf icons to download and open. We recommend using the Word version as it makes formatting the final file easier.



### III. OPEN AUTHORING TEMPLATE

9. The arrangement and formatting of SDG metadata files is not uniform. To standardize (and therefore improve the efficiency and accuracy) of conversion into machine-readable format, we collaborated with the SDMX-SDG Working Group of the IAEG-SDG to prepare an [authoring template](#).

This template has continued to evolve to meet the needs identified by collaborating groups managing SDG metadata, such as the SDMX-SDG Working Group and the IAEG-SDG Secretariat. Some of these versions use a **macro-enabled Word document format** (i.e., **.docm file**). **All of the versions produced, however, are compatible** with the needs of the SDG Translation Project. Below, we note some hints when working with macro-enabled Word files.

SDG Metadata Authoring Template	
Word Version 2.2	
Main and Detailed Concept Collection Form	
Contents	
<a href="#">Metadata Attachment</a> .....	2
<a href="#">Metadata Submission Form</a> .....	2
<a href="#">0. Indicator information</a> .....	2
<a href="#">1. Data reporter</a> .....	2
<a href="#">2. Definition, concepts, and classifications</a> .....	3
<a href="#">3. Data source type and data collection method</a> .....	3
<a href="#">4. Other methodological considerations</a> .....	3
<a href="#">5. Data availability and disaggregation</a> .....	4
<a href="#">6. Comparability/deviation from international standards</a> .....	4
<a href="#">7. References and documentation</a> .....	4
<a href="#">Appendices</a> .....	4

10. This template aligns IAEG-SDG metadata format with the SDMX-SDG [metadata concepts](#). We used the saved template to make easy revisions of text or format.


Metadata Submission Form	
0. Indicator information	
Concept name	<i>Insert text, lists, tables, and images.</i>
0. Indicator information	
0.a. Goal	
0.b. Target	
0.c. Indicator	
0.d. Series	
0.e. Metadata update	
0.f. Related indicators	
0.g. International organisations(s) responsible for global monitoring	

## IV. CUT AND PASTE TEXT ACCORDING TO TEMPLATE MAPPING GUIDE

11. Cut and paste text from the IAEG-SDG file into the template according to the metadata concepts listed on the left-hand side. Many of the metadata concepts will not have corresponding text from the IAEG-SDG file (e.g., series); just leave those blank.

0. Indicator information	
Concept name	<i>Insert text, lists, tables, and images.</i>
0. Indicator information	
0.a. Goal	Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation
0.b. Target	Target 9.1: Develop quality, reliable, sustainable and resilient infrastructure, including regional and trans-border infrastructure, to support economic development and human well-being, with a focus on affordable and equitable access for all
0.c. Indicator	<a href="#">Indicator 9.1.2: Passenger and freight volumes, by mode of transport</a>
0.d. Series	
0.e. Metadata update	Last updated: 13 February 2019
0.f. Related indicators	
0.g. International organisations(s) responsible for global monitoring	<b>Institutional information</b> Organization(s):

12. Some versions of the template (e.g., 3.3) begin with a **series of macro-enabled drop down tables**. These are especially helpful to custodian agencies who are using the template to support transmission of SDG metadata for their indicators, but can also be used by countries submitting their national SDG metadata. If/when prompted, click “**Enable Content**” to select among the drop down tables.

 **SECURITY WARNING** Some active content has been disabled. Click for more details.

Enable Content

### SDG Metadata Authoring Tool Template

Word Version 3.3 (DRAFT)

#### Metadata Attachment

Reporting type  
Reporting Type: Global (G) Reporting Type

SDG series  
SDG Series: 1.1.1 Proportion of population below international poverty line [1.1.1] (SI\_POV\_DAY1) SDG Series

SDG Series: 1.1.1 Proportion of population below international poverty line [1.1.1] (SI\_POV\_DAY1) SDG Series

Add Series Remove Series

Reference area  
Reference Area: World (1) Reference Area

Metadata language  
Metadata Language: en Metadata Language

Please use a 2-letter ISO 639-1 language code. Full list is available at the [ISO web page](#).

[Import Data Structure Definition](#)  
To update the options in the dropdowns according to your SDMX DSD, click the button below:  
Import SDMX DSD

13.

14. Some templates completed for this project were completed using early versions of the template (e.g., version 1). **To include the most recent macro-enabled aspects of the template**, open the most recent template and **cut and paste the entire template starting with “Metadata Submission Form” under the section “Metadata Attachment.”** (You cannot cut and paste the macros into the .docx file, but you can cut and paste from the .docx file into the .docm file.)

SDG Metadata Authoring Tool Template  
Word Version 3.3 (DRAFT)

### Metadata Attachment

---

Reporting type  
Global (G)

SDG series  
1.1.1 Proportion of population below international poverty line [1.1.1] (SI\_POV\_DAY1)

Add Series

Remove Series

Reference area  
World (1)

Metadata language  
en

Please use a 2-letter ISO 639-1 language code. Full list is available at the [ISO web page](#).

#### Import Data Structure Definition

To update the options in the dropdowns according to your SDMX DSD, click the button below:

Import SDMX DSD

From Authoring Tool Test  
16 April 2020

### Metadata Submission Form

---

#### 0. Indicator information

Detailed concept name	<i>Insert text, lists, tables, and images. See Tips.</i>
0.a. Goal	Goal 1: End poverty in all its forms everywhere
0.b. Target	Target 1.1: By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day

Insert early template content (non-macro enabled) here

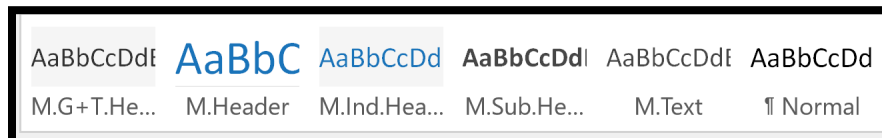
15. For the most part, the order of text is the same between the two documents, but **there are some important differences**. Use the map provided at the end of the template as a guide.

C. Mapping of IAEG-SDG File Format to SDMX Detailed Concepts	
This is an aid to format metadata previously submitted in the IAEG-SDG file format into the format of the SDMX metadata concepts. In some cases, these headers and sections of the IAEG-SDG format do not correspond perfectly to the SDMX concepts. Consult the SDMX metadata concept definitions.	
IAEG-SDG Headers and Subheaders	Detailed Metadata Concept Name
Last updated: [DATE]	0.e. Metadata update
Goal	0.a. Goal
Target	0.b. Target
Indicator	0.c. Indicator
	0.d. Series
Institutional information	0.g. International organisations(s) responsible for global monitoring
Organization(s)	1.a. Organisation
	1.b. Contact person(s)

## V. FORMAT PASTED TEXT USING GUIDANCE

16. Once the text has been pasted into the correct locations in the template, apply the formatting styles as noted below. This ensures that the machine reading conversion retains the appropriate formatting and that a uniform formatting is used to help orient metadata users. (NOTE: Cut and pasted text may **appear** like the actual style has been applied, but it is not always the case.)
17. NOTE: *It is not strictly necessary to use the formats listed below.* Users can use any font to format headers in the metadata files. However, using these formats is desirable. The formats listed below are most consistent with the initial format of the IAEG-SDG metadata files, and applying that style consistently across metadata files will improve readability and therefore user experience. To apply the pre-existing formats, simply select the desired text and click on the style in the Styles bar (under Home).

a. Heading and text formatting



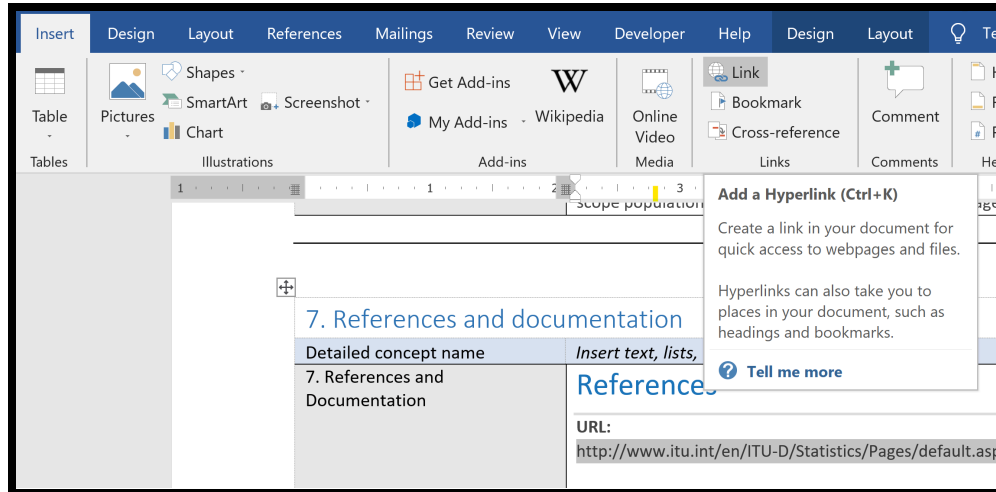
- i. Goal and target: Use M.G+T.Heading
- ii. Indicator: Use M.Ind.Heading
- iii. Last update: Use Calibri (Body) 9 fnt
- iv. Headers (e.g., Institutional information): Use M. Header
- v. Subheaders (e.g., Organization(s):) Use M. SubHeader
- vi. Sub-Subheaders: Use M. Text and apply bold italics.
- vii. Sub-Sub-Subheaders: Use M.Text and apply italics.
- viii. Text: Use M.Text (Calibri 10.5)

b. Bullets

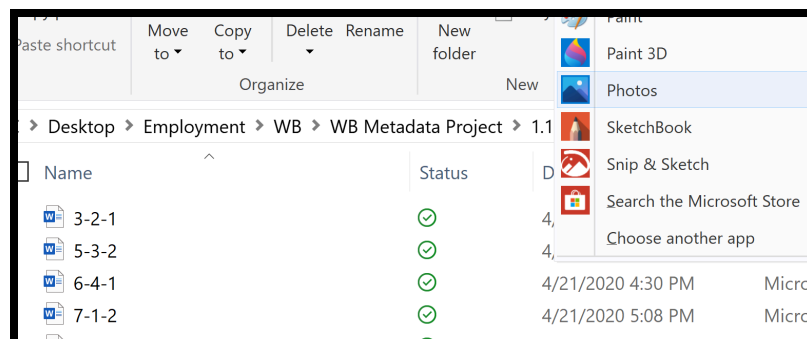


- i. For bullets (\*) use the (first) bullet style.
- ii. For numbered or lettered bullets, use the second bullet style.
- iii. Nested bullets are not supported currently by the software. Nested lists will need to be formatted manually (e.g., 1 a., 1.n, 2.a, 2.b, etc.)

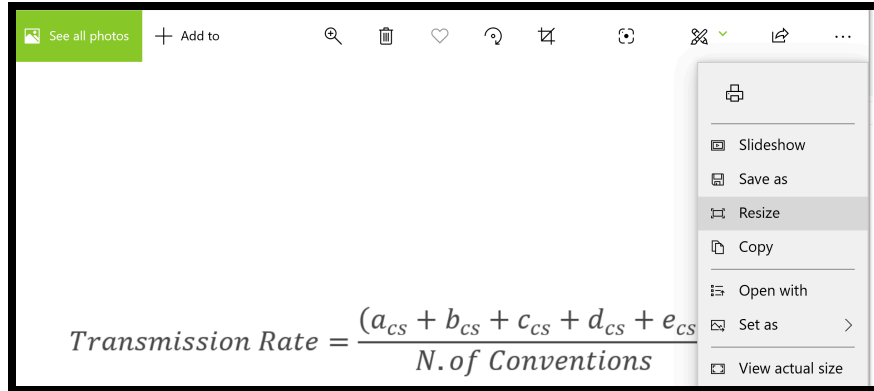
c. URLs: All URLs should have linked hypertext. Choose Insert, Link, and paste the URL to link.



- d. **Footnotes:** Copy and paste from Word document. No additional formatting is needed. Footnotes will be numbered according to appearance in the template, and not necessarily according to appearance on the web page display; however, the footnote text and number will correspond.
- e. **Formulas:** The current authoring template does not work with Word Equation editor. We are exploring a way to resolve this. In the meantime, please treat formulas as images (see below). Formulas typed using usual key strokes convert without difficulty.
- f. **Tables:** Cut and paste from Word document. No additional formatting is needed typically.  
 NOTE: Color is not supported currently by the software. Therefore, if tables use color to indicate values (e.g., progress or data availability), these spaces will appear white when converted. In these cases, the tables will need to be read in as images to retain the color values (see below). Table values using usual key strokes for arrows or checkmarks (e.g.,  $\leftarrow \uparrow \rightarrow \downarrow \checkmark$ ) convert without difficulty.
- g. **Images:** Use Snip application or other method to take a picture of the image to convert it to PNG or JPEG. Copy the file into the template.  
 If the image doesn't fit the width of the table cell, you can reduce it to fit using a editor software. (Some example software tools appear in the screen shot below.) You may need to try out different sizes have it render neatly on the web page.

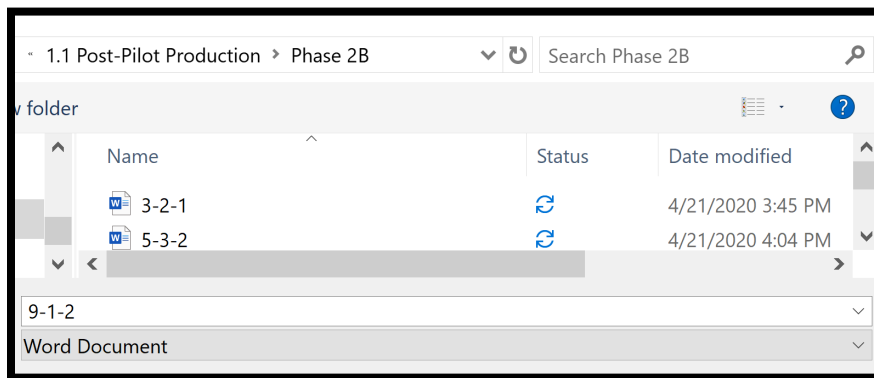






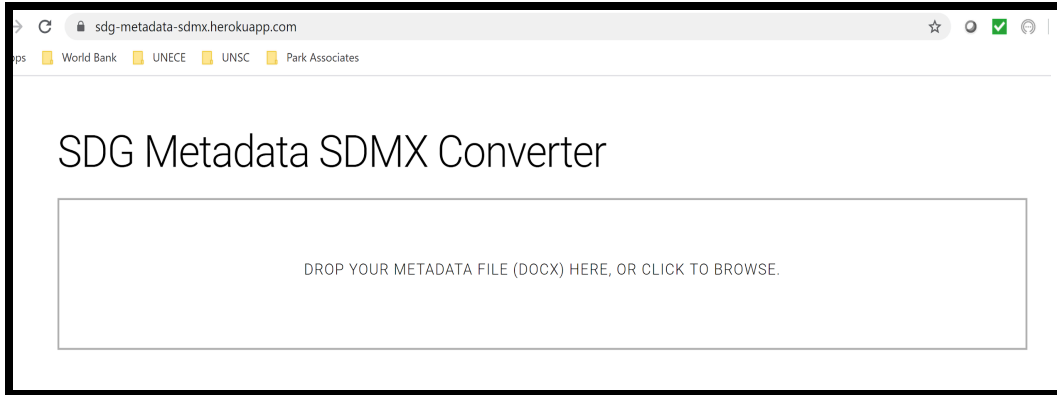
## VI. SAVE AS WORD FILE USING NAMING CONVENTION

18. Each populated template **must** be saved using a specific naming convention. This allows the programs used for conversion into machine readable format to recognize the files. For example, the metadata template file for indicator 3.2.1 must be named “3-2-1.” Use hyphens and no additional text.



## VII. CHECK FORMATTING AS IT WOULD APPEAR IN MACHINE-READABLE FORMAT

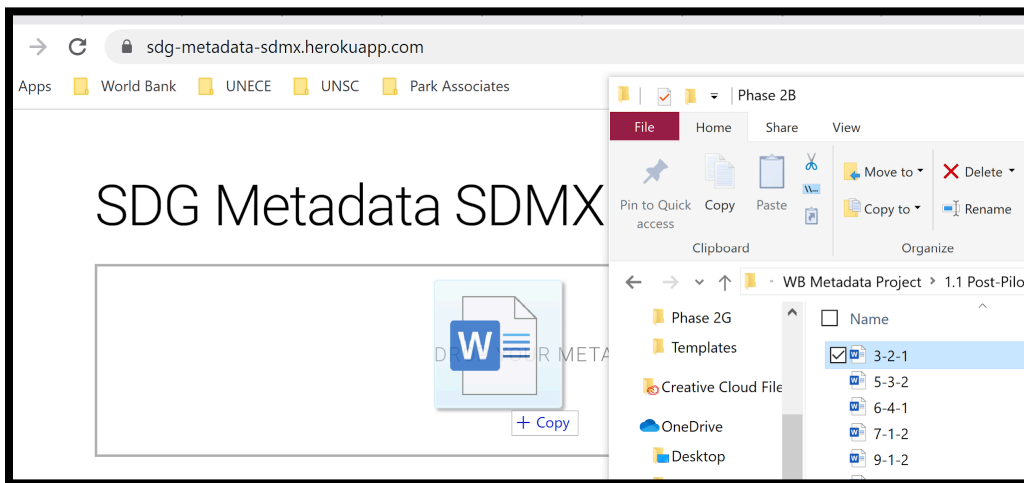
19. As a last check before proceeding to **Step 2: Generate and manage machine readable files in Github for translation in Weblate**, you can check to make sure your formatting is correct before uploading the file onto the SDG Metadata Translation website (and, subsequently, how it would be “read” by Weblate, the translation management service used for this project). **This step is highly recommended**, as it is more difficult to adjust the formatting in the ENG file once in Weblate. In addition, the formatting assigned to the ENG files is automatically assigned to the translated files when completing Step 3.
20. Go to <https://sdg-metadata-sdmx.herokuapp.com/>.



21. The SDG Metadata SDMX Converter is a tool that has two dimensions: 1) it prepares a file of how your formatted machine-readable file will appear on the project website, and 2) it prepares a file in SDG SDMX metadata format.

For the purposes of our project, we only are interested in how the file will appear on the project website, because this will tell us how Weblate will read the file, also.

22. Drag your formatted file into the SDG Metadata SDMX Converter.




23. A zip file is created. Click and open the zip file.

## SDG Metadata SDMX

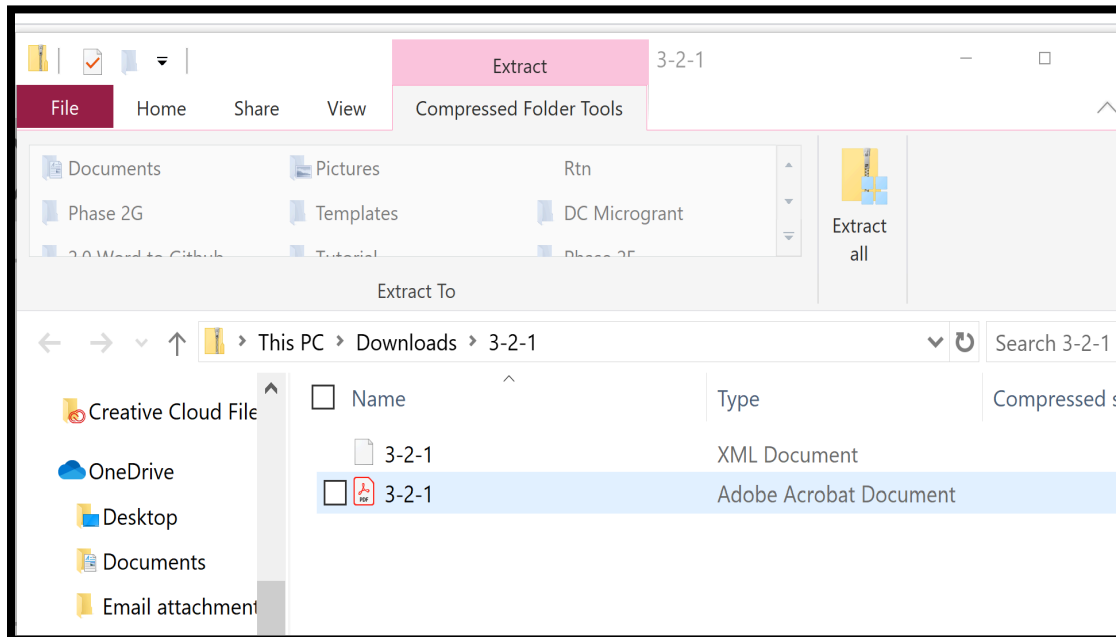
65.1 KB

3-2-1.docx

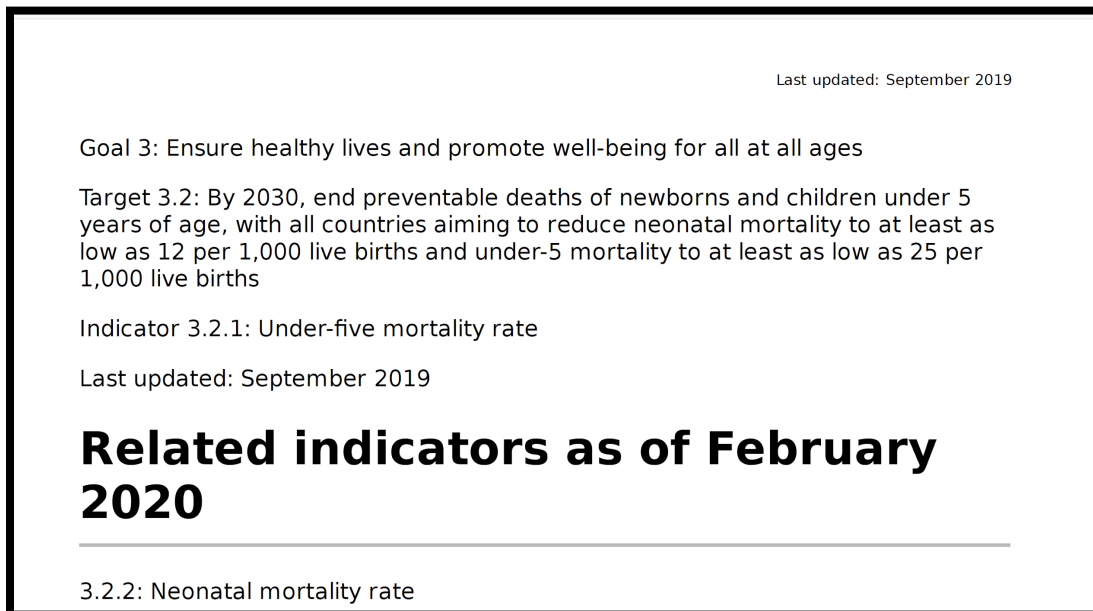
 3-2-1.zip



24. Open the pdf file



25. The pdf file **helps you check formatting** before loading the file into Github and Weblate. These include:, formatting errors in the body text , the headings, the urls, and table and image formats.



26. If needed, **make the formatting changes in the Word template, re-save your file,** and check your fixes by re-loading into the SDMX file converter.

27. **NOTE:** There are **some differences** in how the converter displays the machine-readable file. These differences will not appear when you load them into Github or Weblate. **You can ignore these differences:**
- The pdf file SDG metadata concepts and their associated text are arrayed in SDMX order. This means that they are not in the exact same order as the original IAEG-SDG pdfs.
  - The footnote style and order and order of appearance differs.
  - The date of last update is repeated in the header and in the text.
28. Make any necessary changes in the Word template, re-save the file, and view again using the SDG SDMX metadata file converter.