# 3.2.1 STEP 2.1 - CREATING AND TRANSLATING TEXTS

This document is a step-by-step presentation of how to generate **all the textual content** of the Mooc in any target language. The methodology requires the DeepL API https://www.deepl.com/pro-api?cta=header-pro-api/. It is based on the work carried out by the Inria Learning Lab (Laurence Farhi / Benoit Rospars + the team members primarily involved in the AI4T project: Marie Collin and Bénédicte Cardon) to **produce and distribute the Mooc content** and associated teaching resources in 5 languages.

To illustrate the different stages of translation, we propose a recurring example with the translation of Mooc content into Spanish.

> Note : This document is complemented by a specific-tutorial for the proof reading of the created files once the machine translation step is completed.

## 1- ACCESS TO YOUR FORK OF THE GITHUB REPOSITORY OF THE PROJECT

Example of URL of a created fork from AI4T : **https://github.com/** **\*\*YOURNAME** /ai4t\*\*

## 2- DOWNLOAD YOUR FORK OF THE GITHUB REPOSITORY AS ZIP
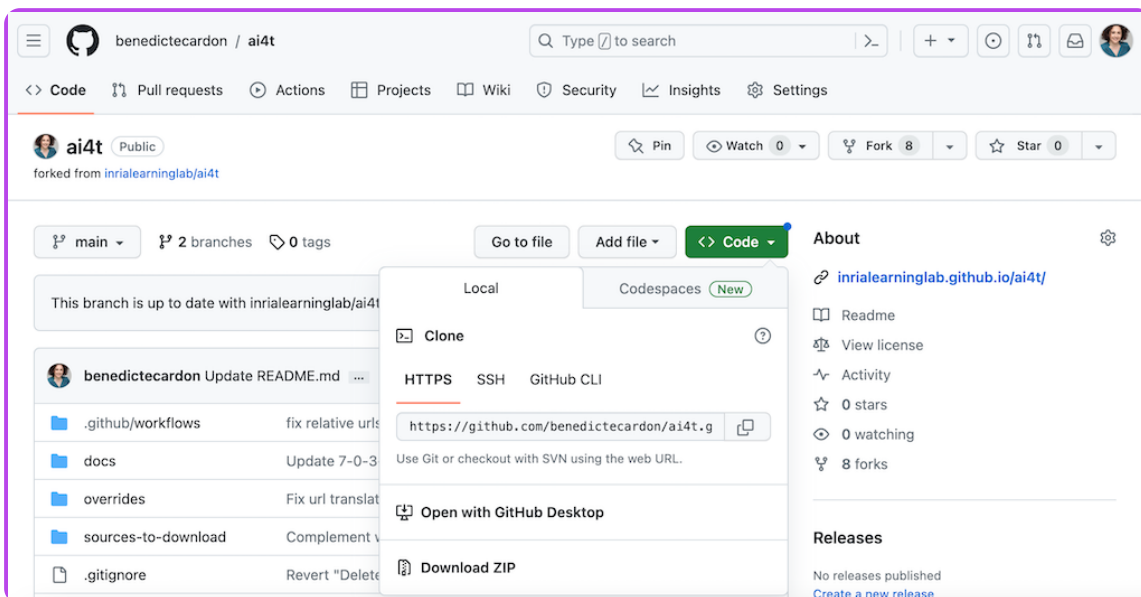


*Downloading a repository in github as ZIP.*

## 3- UNZIP THE FOLDER

You then have a "ai4t-main" folder

## 4- CREATE THE SCRIPT IN YOUR TARGET LANGUAGE

- Copy the file **"trad.py"** and paste it under the relevant name using for example **"tradXX.py"** (🏗 tradES.py for spanish)
- Open the new file **"tradXX.py"** - the code shown before any modification is for translating english files into french.

Then you have to change the target language according to your needs: the **"fr"** mention has to replaced at two places:

```
27                    targetPath = srcPath.replace('.en.md', '.fr.md')
28                    print('Out : ' + targetPath)
29                    if not args.dryrun:
30                        result = str(translator.translate_text(
31                            s,
32                            target_lang="FR"
33                            tag_handling="xml",
34                            ignore_tags=['img', 'a']
```

*Example of extract of the trad.py script with mention of the target language (ie.FR).*

**①** *line 27** : targetPath = srcPath.replace('.en.md','fr.md')*

**②** *and line 32** : target_lang="[FR]",*

## 🏗 For translating to spanish

```
27                    targetPath = srcPath.replace('.en.md', '.es.md')
28                    print('Out : ' + targetPath)
29                    if not args.dryrun:
30                        result = str(translator.translate_text(
31                            s,
32                            target_lang="ES"
33                            tag_handling="xml",
34                            ignore_tags=['img', 'a']
```

*Example of extract of trad.py script with adaptation for translation to Spanish.*

## 5- OPEN A TERMINAL AND ACCESS TO THE "AI4T-MAIN" FOLDER

Run the commande: `cd+space` then drag and drop the "ai4t-main" folder :
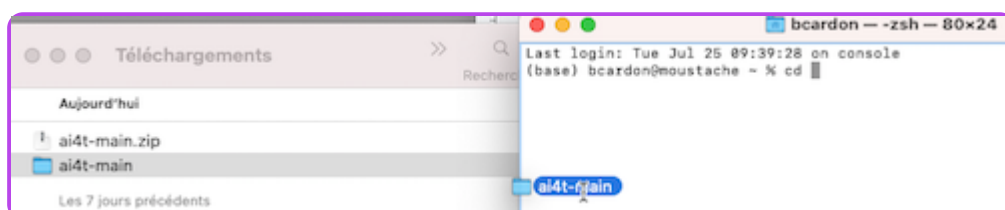
*Visualisation of moving the file from file explorer to terminal.*

## 6- SET UP DEEPL

⌨ Run the command: pip install deepl`

```
% pip install deepl
```

*Code command-pip-install-deepl.*

## 7- GENERATE THE NEW FILES IN TARGET LANGUAGES

⌨ Run the command: `python tradXX.py --key=xxxxxxxx --path=<folder to translate>`

- With **"xxxxxxxx"**: **key** of you DeepL API
- "Folder to translate" : mention a unit folder - use drag and drop

The Mooc is in 4 modules, each module has 3 units (from N-1 to N-3) and an introduction (N-0).

There is also an introductory and a concluding section.

🏗 for translating from EN to ES the folder containing all pages of Module 4 Unit 3 "artificial-intelligence-at-our-service"

```
📁 4-3-artificial-intelligence-at-our-service
```

*Example of unit folder to drag-and-drop for generating translation.*

⌨ Run the command : `python tradES.py --key=xxxxxxxx --path=`

Two new files in Spanish are generated

- 4-3-1v-artificial-intelligence-at-our-service.es.md
- 4-3-2a-case-study-with-ai-templates.es.md

## 8- REVIEW THE CREATED FILES (FORMAL REVIEW)

Using this method, all elements in english in the reference file are translated to the target language.

The some adjustment are to be made for each file as:

- The elements of the header : title, description (and the author and type of resources when relevant)

- And the elements of metadata have to be reviewed to ensure the proposer running of the web pages and the presence of compatible Metadata in each languages.

## 🏗 Translation of a file from EN to ES

```
1    ---
2    title: 1.0.1 Presentation of Module 1
3    description: Presentation of Module 1
4    ---
5    ??? info "Metadata"
6        - Id: EU.AI4T.01.M1.0.1t
7        - Title: 1.0.1 Presentation of Module 1
8        - Type: text
9        - Description: Presentation of Module 1
10        - Subject: Artificial Intelligence for and by Teachers
11        - Authors:
12            - AI4T
13        - Licence: CC BY 4.0
14        - Date: 2022-11-15
15
16    # Presentation of Module 1
```

*Header and Metadata in the the reference EN file.*

```
1    ---
2    título: 1.0.1 Presentación del Módulo 1
3    descripción: Presentación del Módulo 1
4    ---
5    info "Metadatos"
6        - Id: EU.AI4T.01.M1.0.1t
7        - Título: 1.0.1 Presentación del Módulo 1
8        - Tipo: texto
9        - Descripción: Presentación del Módulo 1
10        - Asunto: Inteligencia artificial por y para profesores
11        - Autores:
12            - AI4T
13        - Licencia: CC BY 4.0
14        - Fecha: 2022-11-15
15
```

*Header and Metadata in the the ES file before review markdown file in spanish.*

**Changes to be made:**

**IN THE HEADER**

- **title** instead of **titulo**
- **description** instead of **descripción**

And when relevant:

- **contributors** instead of **colaborador**
- **type** instead of **tipo**
- **text** instead of **texto**
- **activity** instead of **actividad**
- **video** instead **vídeo**

**IN THE METADATA**

- **Title** instead of **Titulo**
- **Tipe** instead of **Tipo**
- **Text** instead of **Texto**
- **Activity** instead of **Actividad**
- **Video** instead **Vídeo**
- **Description** instead of **Descripción**
- **Subject : Artificial intelligence for and by Teachers** instead of **Asunto: Inteligencia artificial por y para profesores**
- **Authors** instead of **Autores**
- **Licence** instead of **Licencia**
- **Date** instead of **Fecha**

**Note:** Be careful with upper and lower case: some elements require a capital letter at the beginning of the word, others do not.
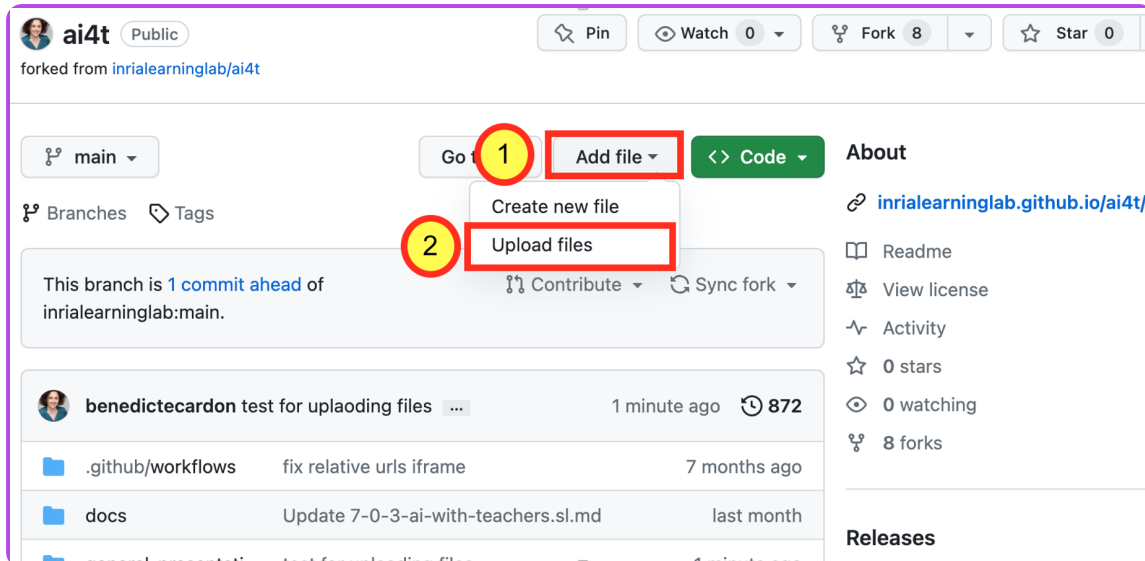


*Header and Metadata in EN and ES files after review markdown files in english and spanish.*

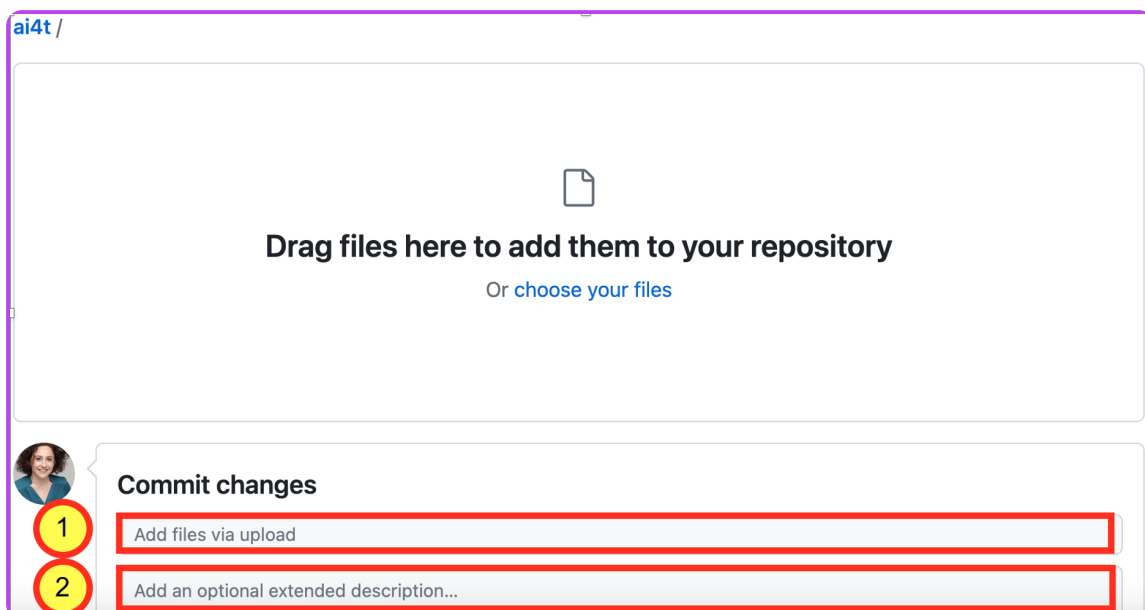## 9- UPLOAD YOUR FILE ON YOUR GITHUB FORK

The files in the new target languages were created locally on your computer. You now need to update your github fork.

**Note** that due to the large number of files, it's not possible to do this in one time.



*Where to upload files in a fork - screen caption of github.*



*How to upload files in a fork - screen caption of github.*

Do not hesitate to add description of your upload for an easier tracking of the fork evolution and notably for synchronisation with other repositories when wanted

## 10- REVIEW THE CREATED FILES (CONTENT + LANGUAGE REVIEW)

Directly on github, in order to:

- Guarantee the linguistic quality of the proposed translation

- Improve or complement the text according to the local context when required

- Suggest substitutions or additions in the target language for resources cited in the original language (articles, books, websites, activities, etc.).

**See dedicated tutorial:** Step 2.2 - Proofreading texts