

# DSpace 7 Translation - Internationalization (i18n) - Localization (l10n)

- [Introduction](#)
- [Getting started](#)
- [2019-11-21 Development status](#)
- [Volunteer!](#)
  - [Dutch \(nl.json5\)](#)
  - [German \(de.json5\)](#)
- [Translator documentation](#)
  - [The authoritative English master file \(en.json5\)](#)
  - [Translations \(for example, nl.json5\)](#)
  - [Starting a new language translation from scratch](#)
  - [Validating your file with the json5 validator](#)
- [Developer i18n How-to](#)
  - [Escaping](#)
  - [Syncing existing translations with changes to en.json5](#)
- [Future work](#)
  - [ICU Expressions and Pluralization](#)
- [What happened to the objective to leverage .po and gettext as new standard?](#)
  - [2019-07-25 State of development \(OLD\)](#)
  - [2019-05-20 State of development \(OLD\)](#)
  - [Grep commands for identifying keys, if we ever want to replace them with the English string again](#)
  - [Background reading](#)

## Introduction

Prior to DSpace 7, the DSpace XML and JSP User interfaces had different catalogs of interface messages. Unified on a single user interface, the DSpace community is transitioning to a single catalog of interfaces messages and better tools for translators.

The DSpace community is actively seeking contributors to aid in the translation of DSpace interface messages, to ensure that DSpace 7 can benefit from the most extensive localization support in the history of the project.

## Getting started

Contributing to translations only requires a (free) Github account. Head over to [www.github.com](http://www.github.com) and get started.

Make yourself familiar with the new json5 catalog of DSpace messages

- <https://github.com/DSpace/dspace-angular/blob/master/resources/i18n/en.json5>
  - This English master file is the authoritative source for the list of keys that are used in the angular interface, as well as the associated English messages
- <https://github.com/DSpace/dspace-angular/blob/master/resources/i18n/nl.json5>
  - As one example of the translated catalog, you will see in the dutch catalog that the English original messages are still present as comments.

The following video walks you through the contribution process:

## 2019-11-21 Development status

Pull Request with placeholder catalogs for all languages we had present in XMLUI: <https://github.com/DSpace/dspace-angular/pull/521>

TODO update the documentation at [Localization L10n](#)

## Volunteer!

Please list your name, email address alongside any of the languages to which you wish to contribute. Also feel free to join the channel [#translation](#) on the [DSpace Slack](#) for assistance and discussion around DSpace 7 translations.

### Dutch (nl.json5)

Bram Luyten - [bram@atmire.com](mailto:bram@atmire.com)

### German (de.json5)

## Translator documentation

### The authoritative English master file (en.json5)

The catalogs can be found in [resources/i18n](#)

#### en.json5 example section

```
"submission.workflow.tasks.claimed.reject.reason.submit": "Reject item",  
"submission.workflow.tasks.claimed.reject.reason.title": "Reason",  
"submission.workflow.tasks.claimed.reject.submit": "Reject",  
"submission.workflow.tasks.claimed.reject_help": "If you have reviewed the item and found it is <strong>not</strong> suitable for inclusion in the collection, select \"Reject\". You will then be asked to enter a message indicating why the item is unsuitable, and whether the submitter should change something and resubmit."
```

### Translations (for example, nl.json5)

In order to allow automated syncing with changes in the English master catalog, and to make it clear to all translators what the original source of the message was, the English original is copied as a comment into the translated files

#### nl.json5 example section

```
// "404.link.home-page": "Take me to the home page",  
"404.link.home-page": "Terug naar de homepagina",  
  
// "404.page-not-found": "page not found",  
"404.page-not-found": "Pagina niet gevonden",  
  
// "admin.registries.bitstream-formats.create.failure.content": "An error occurred while creating the new  
bitstream format.",  
// TODO New key - Add a translation  
"admin.registries.bitstream-formats.create.failure.content": "An error occurred while creating the new  
bitstream format."
```

When new keys are introduced, TODO messages are automatically added to the catalogs of other languages.

### Starting a new language translation from scratch

1. Create an empty file in [resources/i18n](#) and give it the [two letter iso code of your language](#) as the filename, for example `nl.json5` or `fr.json5`
2. Preload your new catalog with all the messages in the `en.json5` catalog by executing the `sync-i18n-files.js` script, for which you can find more detailed documentation lower on this page. Assuming you are in `/resources/i18n`, you can execute:

```
../../scripts/sync-i18n-files.js -s en.json5 -i -t fr.json5
```

where `en.json5` is the "source" language file where the original keys will be retrieved, and `fr.json5` is the catalog of the new language for which you want to start a translation.

The file should now be filled with all of the English messages, that are now present both as comments, as well as message, for you to translate.

3. Translate as many messages as you can. For every message you translate, remove the TODO comment from the catalog.
4. Activate the new language in [environment.default.js](#)

```
{
  code: 'nl',
  label: 'Nederlands',
  active: true,
}
```

Where code is the two letter code for your language, also used as the filename for your translated catalog.

Label is the name of your language, written in that language.

Active set to true, so the language is visible in the language selector on the frontpage.

## Validating your file with the json5 validator

If you have the DSpace 7 UI running locally, including the yarn tooling to start it up, there is a handy command line tool to validate your new translation. For example, to validate the Spanish file "es.json5", you can execute:

### json5 validate

```
yarn json5 --validate ./resources/i18n/es.json5
```

Examples of possible errors you might get:

```
JSON5: invalid character 'R' at 1574:179
```

1574 indicates the line number, and 179 is the 179th character on that line.

This particular mistake occurred where there was an extra space between \ and " in the \" combination used to escape the "

## Developer i18n How-to

When you create new keys, update existing keys or the meaning of existing messages, keep the following in mind.

### Escaping

In the message catalogs, double quotes have to be escaped with \

```
"submission.workflow.tasks.claimed.approve_help": "If you have reviewed the item and it is suitable for inclusion in the collection, select \"Approve\"."
```

## Syncing existing translations with changes to en.json5

Whenever you make one of following changes to en.json5:

- Introduce new keys & messages
- Change an existing key
- Change an existing message

Please run the node script: /scripts/sync-i18n-files.js

The results of your changes will be reflected in the catalogs of the other languages, so translators can pick up the work to:

- Provide translations for the new keys you just introduced
- Update existing translations in case you changed the message for an existing key

By default, the execution of sync-i18n-files.js will:

- look at en.json5 as the authoritative **source** file
- will look at all catalogs present in /resources/i18n/ as **target** files to sync up with the latest changes in source.
- will execute these changes **in-place**, meaning that no backup etc is taken of the translation before the sync

If you want to alter this behaviour, you can:

- specify a different source file using **-s**
- specify a different target file using **-t**
- specify another output file using **-o** in case you want to avoid in place edit

## Future work

### ICU Expressions and Pluralization

Instead of relying on different messages for singular and plural forms, the community hopes to conform the messages to the [ICU Message Format](#) as specified in the [CLDR Plural rules](#)

## What happened to the objective to leverage .po and gettext as new standard?

Originating in the Linux world, [the GNU gettext tools and the PO file format](#) is also the backbone of localization support in [Wordpress](#), [django](#) and [Drupal](#)

The Angular i18n framework we use, NGX translate, has a 3rd party PO file loader: <https://github.com/biesbjerg/ngx-translate-po-http-loader>

Even though the community was initially very optimistic about its potential and the transition to .po and gettext, the major deal breaker was the absence of support for gettext message context (msgctxt), that would allow a translator to translate a key like "Home" into different words in the target language, depending on the context.

The initial ambitions to use the English string as the key itself, and abandon intermediate keys, was also problematic, as we hit a big number of areas in the code where keys were built up programmatically.

As a result, the community settled for:

- JSON5 as the format for the message catalog
- Reverting to a flat list of keys, instead of a hierarchical tree. This now makes it possible again to search on a particular key, which was not possible anymore in the hierarchical format.

### 2019-07-25 State of development (OLD)

The key challenges that are still being tackled are:

- Compatibility between the PO loader for NGX and Angular 7 cfr <https://github.com/biesbjerg/ngx-translate-po-http-loader/issues/29>
- Support for pluralization
- Support for the translation of a specific English message, into different messages in another language, depending on the context.

### 2019-05-20 State of development (OLD)

As part of preview release 1, the developers are still using en.json catalogs. Once [Pull Request 366](#) is accepted, the migration to the new .POT and .PO standard files will be official.

As long as DSpace 7 is still in development, it is expected that the dspace.pot catalog, as well as the different translations, will continue to be extended and evolved.

Together, we aim to release as many, as complete translations as possible, as part of the official DSpace 7.0 release.

## Grep commands for identifying keys, if we ever want to replace them with the English string again

Because we are currently **keeping message keys**, there is no immediate use for identification of keys to replace with the English strings. But for future use, this might still come in handy:

If you execute following command in the angular source directory, you get a list of keys that have not yet been replaced.

#### Grep command for identifying keys that have not yet been replaced

```
grep -snRHIE "'.*\.[^\s]+\.[^']+' \| translate" *
```

Sample output looks like:

```
app/+community-page/delete-community-page/delete-community-page.component.html:5: <h2 id="header" class="border-
bottom pb-2">{{ 'community.delete.head' | translate
app/+community-page/delete-community-page/delete-community-page.component.html:7: <p class="pb-2">{{ 'community.
delete.text' | translate:{ dso: dso.name } }}</p>
app/+community-page/delete-community-page/delete-community-page.component.html:12: <button class="btn btn-
primary" (click)="onCancel(dso)">{{ 'community.delete.cancel' | translate}}
app/+community-page/edit-community-page/edit-community-page.component.html:4: <h2 id="header" class="border-
bottom pb-2">{{ 'community.edit.head' | translate }}</h2>
```

## Background reading

[The PO Format](#)

[GNU gettext utilities](#)