

Dataverse Translation Workshop

Translation of the user interface

Marion Wittenberg
Laura Huis in 't Veld
Veronika Heider

SSHOC Workshop
2nd June 2021



Housekeeping

- Please mute yourself.
- Post your questions in the chat, they will be answered during the Q&A part of the workshop.
- The session will be recorded and made available afterwards.

Program

- ❑ Short introduction about SSHOC task 5.2 by Marion Wittenberg (DANS)
- ❑ Introducing the Weblate tool by Laura Huis in 't Veld (DANS)
- ❑ Experiences from translating to German by Veronika Heider (AUSSDA)
- ❑ Q and A
- ❑ Planning and follow up



Short introduction to SSHOC task 5.2

Marion Wittenberg (DANS)



Project:



SSHOC

social sciences & humanities open cloud



Horizon 2020
European Union Funding
for Research & Innovation

Type of action & funding:
Research and Innovation action
(INFRAEOSC-04-2018)

Partners: 48

(23 beneficiaries + 25 LTPs)

SSH ESFRI Landmarks and Projects
& international SSH data infrastructures

Project budget:

€ 14,455,594.08

Duration: 40 months

(January 2019 – 30 April 2022)

Project website:
www.SSHOpenCloud.eu



Objectives:

- creating the social sciences and humanities (SSH) part of European Open Science Cloud (EOSC)
- maximising **re-use** through **Open Science** and **FAIR** principles (standards, common catalogue, access control, semantic techniques, training)
- interconnecting existing and new infrastructures (clustered cloud infrastructure)
- establishing appropriate **governance model** for SSH-EOSC

SSHOC task 5.2 Hosting and sharing data repositories

- Development of data repository for SSH researchers
- Makes use of open source Dataverse software (IQSS, Harvard)
- Partners: 4 ERICs involved:
 - University of Goettingen, PSNC (DARIAH)
 - UiT (CLARIN)
 - CNR (EHRIS)
 - AUSSDA, DANS (CESSDA)

The main functionalities we are developing

- Kubernetes pipeline (mature software development)
- Development of previewers.
- Development of data migration solutions from legacy archival technologies into Dataverse.
- Translation service for translation of the UI in the European languages
- Connection to localised PID providers
- Integration of Apache Taverna VRE in Dataverse workflow.
- Custom metadata support for Solr and Kubernetes.
- Adjusting the metadata to domain specific needs (CESSDA CMM and CLARIN CMDI metadata)
- Interoperability with domain specific Controlled Vocabularies and Thesauri

Introducing the Weblate tool

- Collaboration and procedures

Laura Huis in 't Veld (DANS)

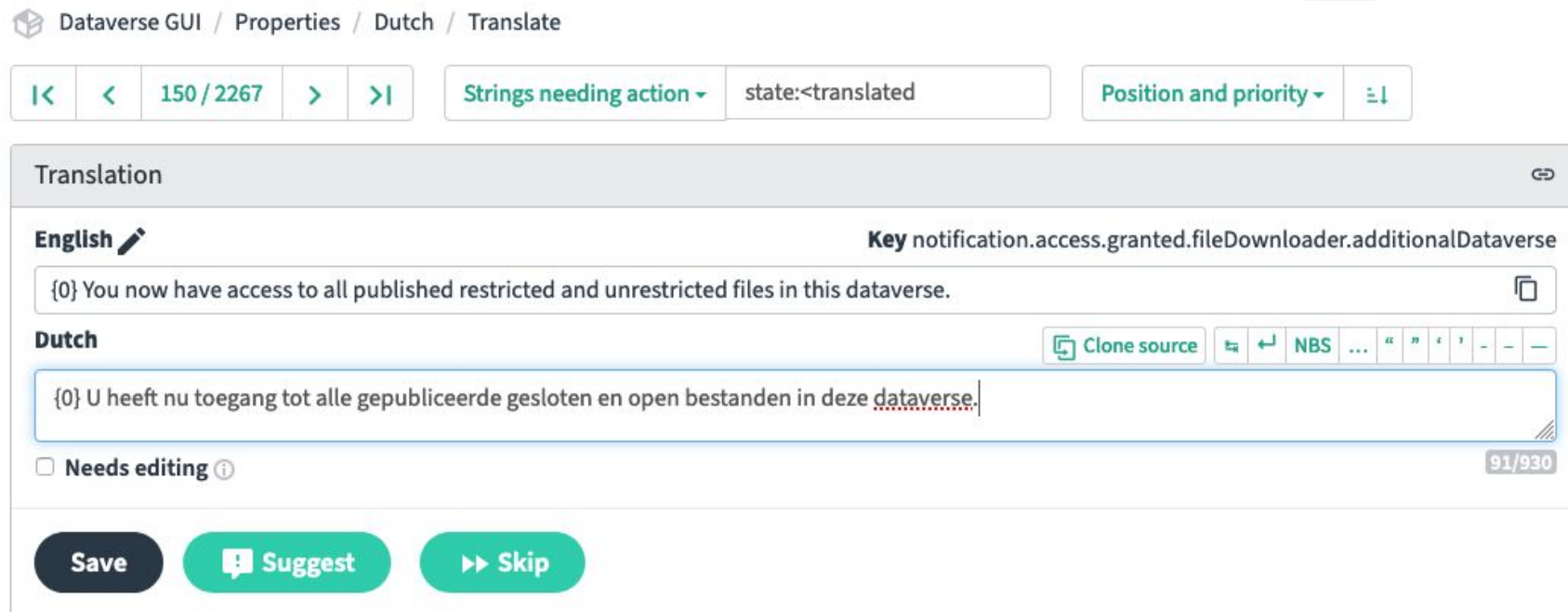


Why Weblate?

- Gives a user friendly interface for translating the dataverse language files;
- Helps you to keep track of your progress.
- When upgrading the dataverse software, you will automatically see the new untranslated strings.
- Collaborate with other translators.

Weblate Basics

- Link for the SSHOC Weblate: <https://weblate-sshoc.cessda.eu>
- Previous Weblate webinar slides¹: <http://doi.org/10.5281/zenodo.4279182>
- See also the Dataverse Translation User guide: <http://doi.org/10.5281/zenodo.4807371>



The screenshot shows the Weblate interface for a translation task. At the top, the breadcrumb navigation reads "Dataverse GUI / Properties / Dutch / Translate". Below this, there are navigation controls including a "Strings needing action" dropdown set to "state:<translated" and a "Position and priority" dropdown. The main content area is titled "Translation" and shows the source string in English: "{0} You now have access to all published restricted and unrestricted files in this dataverse." The target string in Dutch is: "{0} U heeft nu toegang tot alle gepubliceerde gesloten en open bestanden in deze dataverse." The interface includes a "Needs editing" checkbox, a "Clone source" button, and a "91/930" indicator. At the bottom, there are three buttons: "Save", "Suggest", and "Skip".

¹ Wittenberg, Marion, & Huis in 't Veld, Laura. (2020, November). Dataverse, Adaptation to the European Research Environment, Translation of the User Interface (Version 1.0). Presented at the Realising the European Open Science Cloud, online: Zenodo. <http://doi.org/10.5281/zenodo.4279182>

Collaboration within Weblate (1)

Collaboration on language level

- Multiple translators can save suggestions, reviewer(s) can save the final translation.
- Mark translation with 'Needs editing', so others can take a look.
- Make comments or add discussion points for others.
- Add terms in the Glossary.

New comment

Comment on this string for fellow translators and developers to read.

Scope

Translation comment, discussions with other translators

Is your comment specific to this translation or generic for all of them?

New comment

@VeronikaHeider I have also problems with translating this term, do you have any suggestions?

You can use Markdown and mention users by @username.

Save

Collaboration within Weblate (2)

Collaboration between languages

- Link screenshots to source strings for context information.
- Add explanation texts to the English source strings.
- Multi lingual translators can compare translations.
- Make use of already existing translations from previous or later versions of Dataverse.



Weblate Procedures (1)

- As not every participant will have the same version of Dataverse running, a separate project for each version will be created.
 - Glossaries can be shared in between the different projects.
- (Final)Translations can be saved in Github in the appropriate version branche. (Manually for now).
 - <https://github.com/GlobalDataverseCommunityConsortium/dataverse-language-packs/branches>
- Project access control will be set as ‘Protected’: visible for everybody but only chosen users can make changes

Weblate Procedures (2)

- During the SSHOC project, task 5.2 will be responsible for the administrative side of the application. Settings are configured on project level.
- We will enable the review workflow, but it is not mandatory to use reviewers.
- We will enable the possibility to do suggestions.
- Default user settings
 - Translate
 - Screenshots
 - Glossary
 - Languages (temporarily)

Access to the CESSDA Weblate application

URL: <https://weblate-sshoc.CESSDA.eu>

Interested in trying and haven't asked for access yet? Please send an e-mail to: training@CESSDA.eu

If you already have an account, but want to translate an older version of Dataverse, please inform us, so we can create the project within Weblate.



Experiences from translating to German

Veronika Heider (AUSSDA)



Starting point

- First translation to German in a CESSDA ERIC project “DataverseEU”
 - Bundle.properties (Dataverse versions 4.8.6 and 4.9.4)
 - Metadata
 - search facets
- Translation done by a professional translator in Excel

	A	B	D	E
	Original	Translation (Nur der Text nach den = Zeichen ist zu übersetzen; der Text davor soll unverändert stehen bleiben rote Schrift = Übersetzung unsicher oder noch nicht übersetzt!!!)	Comments of the translator	French translation - can be used as a reference
1	(Rot hinterlegte Felder, die mit einem # beginnen, müssen nicht übersetzt werden Text vor den = Zeichen ist nicht zu übersetzen; ebenso wie Text innerhalb dieser Zeichen <>)			
5	dataverses=Dataverses	dataverses=Dataverses		dataverses=Dataverses
6	passwd=Password	passwd=Passwort		passwd=Mot de passe
7	dataset=Dataset	dataset=Datensatz		dataset=Ensemble de données
8	datasets=Datasets	datasets=Datensätze		datasets=Ensembles de données
9	newDataset=New Dataset	newDataset=Neuer Datensatz		newDataset=Nouvel ensemble de données
10	files=Files	files=Dateien		files=Fichiers
11	file=File	file=Datei		file=Fichier
12	restricted=Restricted	restricted=Eingeschränkt		restricted=En accès réservé
13	restrictedaccess=Restricted Access	restrictedaccess=Eingeschränkter Zugriff		restrictedaccess=Accès réservé
14	find=Find	find=Suchen	Auf Deutsch ist "Suchen" der übliche Ausdruck	find=Trouver
15	search=Search	search=Suche		search=Recherche
16	unpublished=Unpublished	unpublished=Unveröffentlicht		unpublished=Non publié
17	cancel=Cancel	cancel=Abbrechen		cancel=Annuler
18	ok=OK	ok=O. K.	O. K. hat auf Deutsch (leider) einen Leerschritt.	
19	saveChanges=Save Changes	saveChanges=Änderungen speichern		saveChanges=Enregistrer les modifications
20	acceptTerms=Accept	acceptTerms=Akzeptieren		acceptTerms=Accepter

Challenges then...

- Translation

- Originally data curators wanted to translate the interface themselves
- Then a translator was hired who had experiences with technological translations

- Gender

- German distinguishes between - for example - male user (Nutzer) and female user (Nutzerin)
- Especially in technological fields, male words dominate
- We decided to use neutral terms where possible
 - “Nutzende” in German, meaning a gender neutral plural form of “users”
 - Supported by German Gender dictionary <https://geschicktgendern.de/>
- We also use medial capital I (*German: Binnen-I*)
 - “ManagerIn” in German, combining male and female forms of “manager”
 - Shorter than “Managerin/Manager”

Challenges then...

- Placeholder / Context

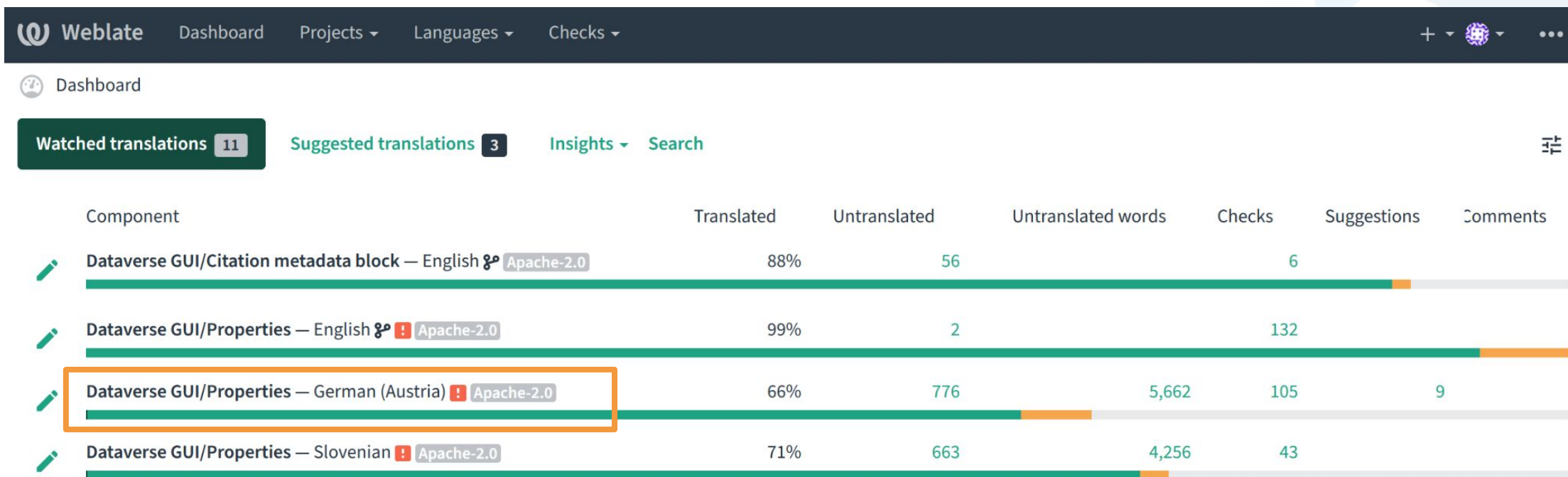
- Hard to tell what the context in the bundle.properties was, for example `{0}` was created in `{1}` . To learn more about what you can do with your dataverse, check out the `{2}`.

Original	Translation	Comments of the translator
notification.createDataverse={0} was created in {1} . To learn more about what you can do with your dataverse, check out the {2}.	notification.createDataverse={0} angelegt in {1} . Was Sie sonst noch mit einem Dataverse machen können, erfahren Sie hier: {2}.	Kommt darauf an, was hinter {1} steht. Im Zweifel: in statt im

Translation in Weblate

- Creating new languages and importing existing files is very easy
- Add context by
 - Adding screenshots to the source language
 - Fill your language's glossary for words that can always be translated the same way
- Cooperate with others
- Use other languages as reference points in the same project

After logging in I decide on the component of my language I want to translate



Strings status				
2,305	All strings — 16,708 words	Browse	Edit	Zen
2	Read only strings — 0 words	Browse	Edit	Zen
1,529	Translated strings — 11,046 words	Browse	Edit	Zen
1,527	Strings waiting for review — 11,046 words	Browse	Edit	Zen
776	Strings needing action — 5,662 words	Browse	Edit	Zen
776	Not translated strings — 5,662 words	Browse	Edit	Zen
9	Strings with suggestions — 34 words	Browse	Edit	Zen
767	Strings needing action without suggestions — 5,628 words	Browse	Edit	Zen
105	Strings with any failing checks — 844 words	Browse	Edit	Zen
105	Translated strings with any failing checks — 844 words	Browse	Edit	Zen
2	Translated strings with dismissed checks — 2 words	Browse	Edit	Zen
58	Failed check: Unchanged translation — 146 words	Browse	Edit	Zen
9	Failed check: Double space — 225 words	Browse	Edit	Zen
29	Failed check: Mismatched full stop — 250 words	Browse	Edit	Zen
1	Failed check: Mismatched question mark — 22 words	Browse	Edit	Zen
1	Failed check: Mismatching line breaks — 11 words	Browse	Edit	Zen
5	Failed check: XML markup — 125 words	Browse	Edit	Zen
3	Failed check: Consecutive duplicated words — 87 words	Browse	Edit	Zen
2,305	Strings without a label — 16,708 words	Browse	Edit	Zen

I then select “Strings needing action without suggestions” to work on “new” untranslated words

Webate Dashboard Projects Languages Checks

Dataverse GUI / Properties / German (Austria) / Translate translated 66%

Strings needing action without suggestions state:<ti Position and priority

Translation English Key contact.context.dataverse.ending

Go to dataverse [2]/dataverse/{3}

You received this email because you have been listed as a contact for the dataverse. If you believe this was an error, please contact [4] at [5]. To respond directly to the individual who sent the message, simply reply to this email.

German (Austria) Clone source

Needs editing 0/2840 · 284

Save Suggest Skip

Nearby strings 31 Nearby keys 30 Comments Automatic suggestions Other languages 4 History

Glossary English German (Austria) Contact Kontakt Dataverse GUI Add term to glossary

String information Screenshot context No screenshot currently associated. Add screenshot Key contact.context.dataverse.ending Flags auto-java-messageformat String age a month ago Source string age 6 months ago

Use the glossary to let Weblate memorize translations of often used words

Enter your translation here (with the “Clone source” function you do not loose breaks)

See other translated languages here



Next up...

- ☐ Q and A
- ☐ Planning and follow up



Q and A

Please add your questions in the chat!



Planning and follow up

- Keep in touch! Please write us your questions to training@cessda.eu
- Want to get started and are interested in a specific version of Dataverse? Get in touch with us!
- After summer, a follow-up workshop will be organised to discuss your user experience.



Thank you for your attention!

Join our community



<https://www.sshopencloud.eu>



@SSHOpenCloud



info@shopencloud.eu



/in/shopencloud

