

Innovation in Terminology Training Through Crowdsourcing: the *Terminology Without Borders* Case

Maria Carmen Staiano¹, Maria Bruno² and Irene Arto Escuerdo³

¹Researcher, Italy

²Researcher, Italy

³ Researcher, Spain

Abstract

This paper presents the Terminology training possibilities offered by *Terminology Without Borders* (TWB), a project of the Terminology Coordination Unit of the European Parliament run through crowdsourcing. TWB provides free multilingual termbases made available to the public on a dedicated website. It can be classified as a Crowdsourcing for Education initiative, as the termbases are created through collaboration among the participants (students, university teachers, professional linguists and subject matter experts), recruited via open calls, and has the goal of educating future terminologists. Another goal of the project is to create a network of universities and individuals interested in Terminology.

Keywords

Terminology training, Terminology Without Borders, crowdsourcing, education

1. Introduction

This paper presents the Terminology training possibilities offered by *Terminology Without Borders* (TWB), a project of the Terminology Coordination Unit of the European Parliament (TermCoord) run through *crowdsourcing*². The following sections focus on the objectives of the project, its innovative aspects, the work done by collaborators, the tools used and the evaluation of the learning outcomes.

2. *Terminology Without Borders* as crowdsourcing

The main goal of TWB is to facilitate communication “by tailoring terminology to citizens’ needs” [3]. This is done by providing multilingual termbases created in collaboration with universities and professional linguists. These termbases are used also to enrich IATE, the terminology database of the European Union institutions [4]. Another goal of the project is to create a network of universities and individuals working in Terminology.

TWB can be classified as a Crowdsourcing for Education (CfE) initiative, “a type of online activity in which an educator, or an educational organization proposes to a group of individuals via a flexible open call to directly help learning or teaching” [5]. Indeed, in line with the requirements illustrated by the same authors:

1st International Conference on “Multilingual digital terminology today. Design, representation formats and management systems”, June 16 – 17, Padova, Italy

EMAIL: carmen-staiano@outlook.it (M.C. Staiano); MBTranslatorAndLinguist@outlook.com (M. Bruno); ireneartoescuerdo@gmail.com (I. Arto Escuerdo)

ORCID: 0000-0003-1458-5225 (M.C. Staiano); 0000-0003-1746-3469 (M. Bruno); 0000-0002-8747-5900 (I. Arto Escuerdo)



© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

² The term crowdsourcing is a portmanteau of “crowd” and “outsourcing” and was coined in 2006 by Jeff Howe in the article on the Wired magazine: “The Rise of Crowdsourcing” [1]. According to Howe, crowdsourcing can be defined as “taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call”. Today, the phenomenon refers to the outsourcing of tasks to groups of individuals carried out through online platforms [2].

- TWB is managed by an organisation that aims to create and share terminological resources for educational purposes.
- Anyone interested in Terminology can contribute to the creation of termbases³.
- The contributors often collaborate on the same termbases.
- One of the purposes of TWB is to train prospective terminologists⁴.

The participation to TWB benefits collaborators both educationally and professionally⁵. Concerning the educational benefits, students and linguists can complement or fine-tune their academic and professional background with hands-on practice in terminology work. Additionally, they acquire new knowledge on specific subjects and learn to use specialised tools. With respect to the professional benefits, they are awarded a certificate of participation, which they can include in a portfolio to showcase their experience. Moreover, all participants become part of a network of students and professionals interested in Terminology. Being part of the community, or the ‘crowd’, helps them to explore future job and educational interests and learn more from their peers.

3. The traditional Terminology training

In many cases, training in Terminology work is given within Translation university courses⁶. Some universities offer targeted Terminology courses for professionals, such as the University of Savoie-Mont Blanc [9] and the University of Barcelona Pompeu Fabra [10]. The University of Trieste has set up a multilingual terminographic database enriched by the work of its Translation and Interpreting students: TERMit⁷. Terminographic records are drawn up under the supervision of the TERMit project managers and, where necessary, of subject matter experts.

Outside universities, terminology training is offered by professional organisations like Termnet, through its yearly Terminology Summer School [11] and the awarding of the ECQA Terminology Manager certificate [12].

We know of another international organisation that, similarly to the European Parliament’s TermCoord, collaborates with university students to enrich their terminology database and train future terminologists: the Terminology Unit of the World Intellectual Property Organisation (WIPO)⁸.

The innovative aspect of TWB, in comparison to the projects mentioned above are:

- TWB is open also to non-students. Usually these are professional linguists that look for additional training or experience in terminology work.
- TWB accepts proposals of terminology projects in any language.
- The work done by all collaborators (students, university teachers, professional linguists and subject matter experts) is promoted and made accessible on a dedicated website.

4. The workflow for the creation of termbases

The four-steps workflow proposed to collaborators for the creation of the termbases is summarised below.

1. Identification of the terms in one or more languages. This step involves term extraction from custom-made corpora and the selection of candidate terms.
2. Compilation of terminological tables.

³ The most frequent collaborators are Terminology students, supervised by their Terminology teachers, and professional linguists (these are subject to a screening to assess their qualifications and experience).

⁴ This is not the main goal of TermCoord, but it is the primary purpose of university teachers.

⁵ These align with the principles of CfE and the general definition of crowdsourcing, as explained by Jiang et al. [5] and by Monti [6].

⁶ For example, this is done at the University of Ottawa, as explained by Bowker and Marshman [7], and at the University of Montréal [8].

⁷ TERMit contains over 30 glossaries available in two or more languages (Italian, English, German, French, Spanish, Croatian, Dutch, Portuguese, Russian, Slovenian) in several specialised fields (law, medicine, economics and others). Additional information on the project is available at <https://lextrain.units.it/?q=termmit>

⁸ The list of universities that participated in this collaboration is on the homepage of WIPO Pearl, the terminology database of WIPO: <https://www.wipo.int/reference/en/wipopearl/>

3. Initial validation of terms and definitions by Terminology teachers, TWB coordinators and domain experts⁹.
4. Final validation by the terminologists of the European Parliament.

Through this workflow, collaborators learn to:

- create corpora,
- extract terminology from texts,
- research definitions for the terms,
- research equivalent terms in different languages,
- exploit terminological resources and
- use terminology management tools and corpus tools.

Universities and individual linguists can choose to work on the entire workflow and create a new subproject or to work on one or two of the steps. In the latter case, they can select one of the termbases already compiled and find equivalents in their working languages.

5. The tools used

The main tools proposed to collaborators for the creation of corpora and the extraction of terms are Sketch Engine [13] and AntConc [14]. The former is free for all students at EU universities until March 2022, while the latter can be downloaded for free by anyone.

Since March 2021, collaborators have been encouraged to use of the FAIRterm application [15] for compiling the terminology records, as opposed to an Excel table. FAIRterm is a free tool built according to the FAIR principles (Findability, Accessibility, Interoperability and Reusability). It was created as a guideline for those who want to enhance the ability of machines in finding and choosing the data and to support its reusability [16].

Integrating FAIRterm in the TWB workflow helped to standardize and harmonise the termbases created by the collaborators. Indeed, collaborators are not given the possibility to edit the data fields and cannot save their entries without filling in the compulsory fields, which enables the creation of termbases consistent with each other.

6. Evaluation of the learning outcomes

In order to collect feedback and evaluate the collaborators' learning outcomes, an online survey¹⁰ was sent to all participants of TWB. The results will provide insights on how to better adjust the workflow of TWB to address the collaborators' educational needs.

The work done by the collaborators is evaluated either by university teachers or TWB's coordinators¹¹. The Steering Committee of TWB, composed of some of the university teachers involved in the project, has proposed the creation of a pool of subject matter experts who could revise the terminological entries. This would provide additional feedback to collaborators and improve the quality of their final work.

7. Conclusion

This paper presented *Terminology Without Borders*, a Crowdsourcing for Education initiative coordinated by TermCoord. Through voluntary collaborators, TWB provides multilingual termbases freely accessible on yourterm.eu. By working on the project, collaborators learn to perform terminology work, become part of a network of people interested in Terminology and gain recognition for their work. At present, to improve the quality of the termbases, TWB is in need of more experienced

⁹ The initial validation is done by one or more of these actors, depending on their availability.

¹⁰ The survey was sent on 7th February 2022 and will run throughout the following months to enable all participants to answer the questionnaire at their own pace, but also to gather a significant amount of data.

¹¹ Provided that they have professional knowledge of the languages of the entries.

terminologists and subject matter experts for the validation phase for both EU and non-EU languages¹². The recruitment of these experts could be done within the partner universities or through crowdsourcing platforms. In addition, it is essential to increase the involvement of the ‘crowd’ by sending periodic surveys and by creating focus groups. Gathering feedback from collaborators will help TWB coordinators to develop new and more personalised learning journeys, improve the training material¹³ and provide the necessary skills to be a terminologist.

8. References

- [1] J. Howe, The Rise of Crowdsourcing, 2006. URL: <https://www.wired.com/2006/06/crowds>
- [2] J.H. Kietzmann, Crowdsourcing: A revised definition and introduction to new research, *Business Horizons* 60, volume 2 (2017) 151–153. <https://doi.org/10.1016/j.bushor.2016.10.001>
- [3] Terminology Without Borders, Homepage, 2021, URL: <https://yourterm.eu/>.
- [4] IATE European Union Terminology, IATE User’s Handbook, 2021, URL: https://iate.europa.eu/assets/IATE_Handbook_public.pdf
- [5] Y. Jiang, D. Schlagwein and B. Benatallah, A Review on Crowdsourcing for Education: State of the Art of Literature and Practice, in: M. Tanabu and D. Senoo (Eds.), *Proceedings of the 22nd Pacific Asia Conference on Information Systems (PACIS 2018)*, volume 180, 2018. URL: <https://aisel.aisnet.org/pacis2018/180/>
- [6] J. Monti, Crowdsourcing e Cloud-computing: una nuova generazione di risorse e tecnologie per la traduzione, in: C. Bosisio and S. Cavagnoli (Eds.), *Comunicare le discipline attraverso le lingue: prospettive traduttiva, didattica, socioculturale*, Guerra Edizioni, Perugia, 2013, pp. 85-100.
- [7] L. Bowker and E. Marshman. “Better integration for better preparation: Bringing terminology and technology more fully into translator training using the CERTT approach.” *Terminology* 15.1 (2009): 60-87.
- [8] Université de Montréal, Baccalauréat en Traduction, 2021. URL: <https://ling-trad.umontreal.ca/programmes-cours/programmes-traduction/baccalaureat-en-traduction-1er-cycle/>
- [9] Université Savoie Mont Blanc, University Diploma Foundations and Issues of Terminology, 2018. URL: <https://www.univ-smb.fr/formation-continue/en/diplome-universite-terminologie-iufc/>
- [10] Universitat Pompeu Fabra, Máster Online en Terminología, 2021. URL: <https://www.upf.edu/web/terminologiaonline>
- [11] TSS 2021 – TermNet’s ONLINE Terminology Summer School from 1-4 July 2021, 2017. URL: <https://termnet.eu/tss2021>
- [12] TermNet International Network for Terminology, ECQA Certified Terminology Manager, 2018. URL: http://www.termnet.org/english/products_service/ecqa_ctm/ecqa_ctm.php
- [13] M. Jakubíček, A. Kilgarriff, V. Kovář, P. Rychlý and V. Suchomel, Finding Terms in Corpora for Many Languages with the Sketch Engine, in: S. Wintner, M. Tadić and B. Babych (Eds.), *Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, Gothenburg, Sweden, 2014, pp. 53-56. doi: 10.3115/v1/E14-2014
- [14] L. Anthony, AntConc (Version 4.0.4) [Computer Software], 2022. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- [15] F. Vezzani. “La ressource FAIRterm : entre pratique pédagogique et professionnalisation en traduction spécialisée.” *Synergies Italie* 17 (2021): pp. 51-64.

¹² Since the official languages of the EU are 24, there are 24 Language Units within the Directorate for Translation of the European Parliament, and therefore only terminologists with professional competence in EU languages.

¹³ Similarly, the NGO Translators Without Borders provides its volunteers with training on its translation environment and toolkits, besides having a Code of Conduct for translators, where it explicits the standards for a “good” translation [17].

- [16] M. Wilkinson, M. Dumontier, I. Aalbersberg et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3 (2016). doi: <https://doi.org/10.1038/sdata.2016.18>
- [17] N. Krinat. "The challenge of quality management in crowdsourced translation: the case of the NGO Translators Without Borders." *QScience Connect Special Issue-The 10th International Translation Conference 2021.3* (2021). doi: <https://doi.org/10.5339/connect.2021.tii.4>

9. Appendices

Appendix 1

2022 TERMINOLOGY WITHOUT BORDERS SURVEY

This survey aims at collecting feedback on the participation to "Terminology Without Borders" for research purposes. Please note that all your data (i.e. contact details) will be treated confidentially.

1. E-mail*

2. What best defines your role? *
 - Student
 - Professional linguist (translator, interpreter, terminologist)
 - Other: _____
3. What did you expect to achieve by participating in "Terminology Without Borders"? (more than one answer possible)
 - Being included in a network of people interested in terminology
 - Acquire hands-on practice in terminology work
 - Learn more about a domain of knowledge (such as medicine or finance)
 - Obtain recognition of your work (certificate of participation, promotion on yourterm.org)
 - Other: _____
4. Did you have any previous experience in terminology work? *
 - Yes
 - No
5. If you answered "Yes", where did you gain experience in terminology work?
 - University course
 - Online course
 - Training on the job
 - Other: _____
6. How useful was the training material provided by the Terminology Without Borders coordinators (PowerPoint, Excel sheets, live workshops)? *

	1	2	3	4	5	
Not at all helpful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Extremely helpful
7. What did you find most challenging while working on "Terminology Without Borders"? *
 - Using the tools

- Doing terminology work (creating corpora, extracting terms, researching equivalents, etc)
- Collaborating with other participants
- Communicating with the coordinators of the project
- Other: _____

8. How satisfied are you with your experience in "Terminology Without Borders"? *

	1	2	3	4	5	
Dissatisfied	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very satisfied

9. What did you achieve by participating in "Terminology Without Borders"? (more than one answer possible) *

- I learnt more about Terminology theory and practice
- I learnt to use some tools employed in terminology work
- I learnt more about one or more domains of knowledge
- I explored future job/educational interests
- I networked with other people interested in Terminology, Translation and Interpreting
- I did not learn anything new
- Other: _____

10. What did you like the most about the project?

11. What did you like the least about the project?
