

# **COST Action**

## **Progress Review at 24 months**

### **(07/03/2017 to 07/03/2019)**

# **CA16105: European Network for Combining Language Learning with Crowdsourcing Techniques**

This report is submitted by the Action Rapporteur in fulfilment of the requirements of the rules for COST Action Management, Monitoring and Final Assessment and is confidential to the COST Association and the Management Committee of the Action.

## Summaries

### **The main aim and objective of the Action is to**

The enetCollect Action aims at setting into motion a Research and Innovation trend for combining Language Learning and Crowdsourcing in order to enhance the production of language learning material and language-related datasets.

### **During its first two years the Action progressed the achievement of this as described below**

During the first two years of its lifetime the enetCollect Action has progressed well toward its main objective of setting an R&I trend in combining Language Learning and Crowdsourcing into motion. The Action, even though starting from very limited research connections, has managed to build a large, interdisciplinary, active and forward-thinking community of relevant stakeholders. Knowledge about and practical experiences on the subject have been shared among its members and several efforts in extending, implementing and disseminating them further have been started.

The "MoU Objectives" and "Deliverables" have been addressed to a substantial extent and enough traction has been created to be able to catch up on the parts that need some more investment. Indeed, the state-of-the-art reviews and knowledge creation (MoU 1) still needs more systematic organization, structuring of the knowledge and codification into comprehensive documentation (i.e. planned with the final Handbook), while the work on prototypes, dissemination, community building, communication means, additional funding and sustainability (MoU 2 to 7) is either at or beyond the point of development expected. The same favorable conclusion can be drawn for the deliverables: whereas some are behind schedule (6), most of them are progressing as expected (29) (while 3 have been dropped after reconsidering their usefulness or adequacy).

In regard to "Additional Outputs", enetCollect members have proven to be quite prolific in terms of publications and project submissions, in particular when considering the highly novel nature of the effort to combine Language Learning and Crowdsourcing for the purpose of mass-producing language learning material, and language-related datasets at the same time. Other types of output, including working prototypes, videos and bibliographical collections have also been delivered.

In relation to enetCollect's foreseen "Impacts", within the first two years of the Action enetCollect has successfully laid the groundwork for making the impacts described in its MoU a reality. First and foremost, the Action has succeeded in creating a new R&I community by involving stakeholders from various relevant and related communities. By now, the community is involving a large number of stakeholders and is balanced in terms of gender, seniority and nationality. At the same time, the efforts of enetCollect's members in creating the first ideas and prototypes have paved the way towards the mass-production of language learning material and language-related datasets for all languages learned nowadays and thus the way towards the scientific, technological, and socio-economic impacts they directly or indirectly imply.

With reference to "Dissemination" the Action can report on the successful creation of various relevant communication channels, 10 scientific presentations in the context of STSM stays, 4 project presentations at relevant stakeholder conferences, as well as more than 10 written (and audio) communications to the general public or wider scientific audiences which explain enetCollect's overall ambition and give visibility to the Action.

All in all, the Action is progressing smoothly for most parts and is gaining a traction that renders its next achievements even more promising.

### **The Rapporteur summarised the Action's major outcomes, impacts and successes as follows**

During the first two years of the enetCollect Action project members of the network managed to combine Language Learning and Crowdsourcing in order to produce language learning material and language-related datasets in line with their major objective: Research and Innovation. They have managed to create a large, interdisciplinary network. They have addressed the MoU objectives and Deliverables and some are in progress or foreseen in the next two years of the project. Regarding Additional output, enetCollect members have numerous publications and project submissions. EnetCollect Cost Action has societal, technological, economic and scientific impact on the society in the area of Research and Innovation. The network is balanced in terms of gender, seniority and nationality. It has rich output and dissemination, publications, presentations, STSMs, communication to the general public that promote the visibility of the Action. Overall, the Action has a smooth progression to the next phase of the project.

### **The Action Chair has described their plans for addressing issues identified in their report as follows**

All the potential endangerment we foresee is members-related.

[ CG ] In order to address some underperforming of the CG, the Chairs fostered a progressive change in its members by substituting 5 members, introducing 4 new ones and reorganized 2 others in their role. Some members still suffer from limited time availability, which the Chairs help with by holding, when relevant, regular and recurrent meetings to accompany them in the tasks to come and find people to support them in subtasks.

[ Chair ] The Chairs will keep on addressing the "Chair bottleneck" by reorganizing the CG, completing milestones (e.g. first version of the website) and defining procedures for regular tasks, while delegating their implementations (e.g. STSM calls).

[ Members ] To improve how funding is allocated to members, the CG will take further measures to reduce the expenses of COST-intensive members and of passive members, such as formulating guidelines for saving money and elegantly/softly encouraging their application (e.g. booking flats instead of hotels for longer STSM stays, making early travel arrangements) and raising awareness of average expenses, involvement etc.. Also, the Chairs will stop inviting known passive MC members for the entire Annual Meetings if they do not contribute outside of the MC meetings.

### **The Rapporteur summarised the Actions's plans for addressing issues identified in the report**

The plans described above and recommendations are adequate and can maximise the chances of the Action achieving its MoU Objectives.

It was decided to substitute 5 members and to introduce 4 new ones, while the roles of the other 2 members were reorganised.

One of the major problems is limited time availability of the Cost Action members. The Chairs tried to address the issue, they tried to accompany them in the tasks, to have more frequent meetings, to find more people who can help with the subtasks.

The Chairs will continue the reorganisation of the CG in order to manage the completion of the target milestones, regular tasks and procedures.

The procedure of the allocation of funding to active and passive Cost Action members is going to be improved in order to reduce expenses and increase efficiency and research output via raising awareness about expenses and involvement of active researchers and experts and not inviting passive MC members.

Overall, these Action plans and recommendations seem to be relevant for addressing issues identified in their report.

## Achievement of MoU objectives, deliverables and additional outputs/ achievements

### MoU objectives

The Action reported progress of the following objectives.

MoU objective	Level of progress	Rapporteur assessment
Creating a theoretical framework for achieving a shared understanding of the challenge, ranging from very concrete aspects such as defining a common terminology, to higher-level aspects such as evaluating the potential and limits of related approaches.	26 - 50%	CONFIRMED
Gathering evaluation data for complementing the theoretical framework with factual information obtained from prototypical experiments and wide-ranging surveys in order to extend and root the theoretical framework and evaluate which aspects should be approached first.	26 - 50%	CONFIRMED
Disseminating the knowledge created and the achievements to any relevant audience.	51 - 75%	CONFIRMED
Creating a balanced interdisciplinary core community of active stakeholders.	76 - 100%	CONFIRMED
Establishing communication channels allowing to easily share information and reach relevant members or stakeholders.	51 - 75%	CONFIRMED
Fostering new funded initiatives complementing the Action.	51 - 75%	CONFIRMED
Creating an association continuing the work started in the Action after its conclusion.	0 - 25%	CONFIRMED

### Rapporteur assessment of the level of progress reported by the Action.

The Rapporteur confirmed the level of progress reported by the Action for all objectives.

### Action explanation regarding MoU objectives reported as 25% or less achieved

The table below shows the Action's explanation and the Rapporteur's analysis thereof for any MoU objectives that the Action reported as 25% or less achieved.

MoU Objective that was reported as 25% or less achieved	Action's explanation	Rapporteur's analysis
Creating an association continuing the work started in the Action after its conclusion.	The work with respect to this objective has yet to be started since it will be addressed with Deliv. 37 and 38 which are due at the very end of the Action (February 2021).	The Action Chair's explanations are valid as The work with respect to this objective has yet to be started since it will be addressed with Deliv. 37 and 38 which are due at the very end of the Action (February 2021).

### General Assessment of MoU objectives

The level of ambition of the MoU objectives is **Low**.  
Overall, **most MoU Objectives are progressing appropriately**.

## Deliverables

The table below shows, for each deliverable, the delivery status reported by the Action and the Rapporteur's comment.

Deliverable	Month deliverable due	Delivery status	Rapporteur Comment
D.1.1 - Summary report on existing material and the workflows for producing it	12	Not foreseen	This Deliverable is not expected to be delivered due to valid reason, in particular because of the nature of the Cost Action members, linguists, language teachers and no publishers.
D.1.2 - Summary report on the state of the art of explicit crowdsourcing approaches	12	Not delivered, but expected before end of Action	The plans described by the Action are sufficient to ensure the delivery
D.1.3 - Summary report on the theoretical framework for producing material through explicit crowdsourcing	24	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.1.4 - Design plans, implemented prototypes and related evaluation reports on explicit crowdsourcing for language learning material production	45	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.2.1 - Summary report on existing material and the workflows for producing it	12	Not foreseen	The Action Chair's explanation is valid and it is due to the nature of the Cost Action members, linguists and language teachers and no publishers.
D.2.2 - Summary report on the state of the art of implicit crowdsourcing approaches	12	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.2.3 - Summary report on the theoretical framework for producing material through implicit crowdsourcing	24	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.2.4 - Design plans, implemented prototypes and related evaluation reports on implicit crowdsourcing for language learning material production	45	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.3.1 - Summary report on existing state of the art solutions for user-oriented design strategies	12	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.3.2 - Summary report on the theoretical framework with regards to design strategies for functionalities, interface and usability	24	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.3.3 - Design plans and evaluation reports on implemented prototypes	45	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.4.1a - Technology-oriented specifications and blueprints to support explicit and implicit crowdsourcing for language learning material production (WG1 and WG2)	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.4.1b - Expert evaluation of specifications and blueprints to support explicit and implicit	45	Not delivered, but expected before	The plans described by the action are sufficient to ensure the

crowdsourcing for language learning material production (WG1 and WG2)		end of Action	delivery.
D.4.2a - Technology-oriented specifications and blueprints to support user-oriented design strategies for a competitive solution (WG3)	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.4.2b - Expert evaluation of specifications and blueprints to support user-oriented design strategies for a competitive solution (WG3)	45	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.5.1a - Ethical guidelines, including templates of relevant information sheets and forms	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.5.1b - Expert evaluation of the ethical guidelines, including templates of relevant information sheets and forms	45	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.5.2a - Legal guidelines, including templates of relevant information sheets and regulations	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.5.2b - Expert evaluation of the legal guidelines, including templates of relevant information sheets and regulations	45	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.5.3a - Business guidelines, including templates of relevant business model descriptions	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.5.3b - Expert evaluation of the business guidelines, including templates of relevant business model descriptions	45	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.6.4 - Description of the IPR regulations and the means to implement them	6	Not delivered, but expected before end of Action	After inquiring among enetCollect members, for the moment it appeared that members from other language-related R&I fields had also no or little experience in such questions, which in turn indicated that IPR-related issues are unlikely to occur. This question was thus put on hold. It is not clear whether the delivery is going to take place before the end of Action.
D.6.5 - Guidelines for the management and implementation of the Collective Workforce	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.7.1 - Guidelines for the management and implementation of the Dissemination Plan	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.7.2 - Guidelines for the management and implementation of the Exploitation Plan	36	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.7.3 - Guidelines for the management and implementation of the Outreach Plan	36	Not foreseen	This Deliverable is not expected to be delivered as it turned out that the network is growing organically at a steady pace and a formal Outreach Plan seems not needed.

D.7.4 - Contacts database to reach relevant institutions, associations and individuals	12	Not delivered, but expected before end of Action	The plans described by the action are sufficient to ensure the delivery.
D.7.5 - Communication means including project website and social media	12	Delivered	

## General Assessment of deliverables

The level of ambition of the deliverables is **medium**.  
Overall, **some deliverables are progressing appropriately**.

## Additional outputs / achievements

### Co-authored Action publications

The Action reported 9 publications on the topic of the Action, co-authored by at least two Action participants from two countries participating in the Action. The full list of publications appears in Annex I.

The:

- **Quality** of the Action's co-authored publications is **very good**.  
Overall, the enetCollect Cost Action managed to publish various papers that are of high quality and high relevance to the research projects, aims and objectives of the action and its theoretical framework. Overall, the publications present an overview of the enetCollect project, its aims, research trends with the focus on combining Language Learning and Crowdsourcing approaches, its interdisciplinary nature and potentials of crowdsourcing for language learning. They describe the objectives of the Cost Action, its structure: WG1: Explicit Crowdsourcing, WG2 Implicit Crowdsourcing, WG3 User-oriented design strategies, WG4 Technology-oriented specifications for a flexible and robust solution, WG5 Application, specification for an ethical, legal and profitable solution. They also deal with dissemination, exploitation and outreach coordination, relationship between enetCollect and Natural Language Processing (NLP) and related work. They provide a very good overview of the Cost Action that helps to promote its visibility.
- **Significance** of the Action's co-authored publications is **very good**.  
All the publications that have been produced by the enetCollect Cost Action are of great significance to the Cost Action and wider audience. The papers aim to introduce enetCollect project and present its objectives and approach, its relevance to Computer Assisted Language Learning (CALL), structure of the network, implementation strategies and future work. Such publications are of great importance as they promote the work of the Cost Action and all stakeholders involved, present its interdisciplinary framework, its terminology, key concepts, objectives and opportunities. It allows to involve new participants and collaborators, teachers and educators as well as CALL experts, to find new research ideas and initiatives.
- **Relevance** to the Action of the Action's co-authored publications is **very good**.  
All the publications that have been produced so far by the enetCollect Cost Action, its members and collaborators, are highly relevant to the Cost Action, as they provide an overview of the enetCollect Cost Action, its aims and objectives, structure, theoretical framework and practical application of the knowledge constructed within the enetCollect Cost Action. One of such papers deals with DialettiBot, a Telegram based chatbot for crowdsourcing geo-referenced voice recordings of Italian dialect. It aims to collect voice recordings and deals with linguistic variation and preservation of local Italian dialect (proverbs and idioms) and contributes to several voice-based Natural Language Processing (NLP) applications. It helps the reader to understand better non-standard Italian dialects. This paper is a good example of one of the practical applications of knowledge constructed within enetCollect project. It gives an overview of the Italian dialect and related work and projects on Italian dialects. It describes the system of Telegram Chatbot for Crowdsourcing purposes, its technical specification, collected data of the project, major findings and future work.
- **Quantity** of the Action's co-authored publications is **very good**.  
The members of the enetCollect Cost Action have shown the evidence of high quality publications. The number of the publications so far within the period of 24 months could be considered very good as it is very time consuming in order to prepare a publication that is of high quality, relevance and significance to the Cost Action. This is quite an achievement. The participants of the enetCollect managed to do this. Most of the journal articles, conference proceedings or conference presentations deal with a general introduction of the Cost Action, its aims and objectives, its structure, working groups, theoretical framework and practical implementation. Some of the papers focus on the domain of Language Learning and Crowdsourcing approaches in order to enhance the production of language learning materials, build international and interdisciplinary research and innovation community, to create a comprehensive theoretical framework and run prototypical

experiments. One of the papers deals with edutainment, gamification, Computer Assisted Language Learning (CALL) and individual customization. The paper describes SUSTAIN-ALL, a pan-European network of story-based distributed crowdsourcing intelligent knowledge base for language learning and its implementation as well as major findings and future research.

## Projects and proposals resulting from Action activities

The Action reported the following projects resulting from Action activities involving at least one Action participant.

Title	Main proposer name	Funder
WIKOLLECT (resubmission)	Alberto BARRÓN CEDEÑO	H2020
CALLECTOR	Emmanuel Rayner	National
Multimodal Child Acquisition	Costanza Navarretta	National
Large-Scale Information Extraction and Gamification for Crowdsourced Language Learning	Pavel Smrz	National
From parallel corpora to multilingual exercises - Making use of large text collections and crowdsourcing techniques for innovative autonomous language learning applications	Johannes Graën	Other

In addition the Action reported 3 proposals resulting from Action activities involving at least one Action participant, and for which the Action networking was necessary.

Relevance of the Action's proposals and/ or projects is **excellent**.

Quantity of the Action's proposals and/ or projects is **excellent**.

## Other outputs / achievements

The table below shows the other outputs / achievements.

Other Output / Achievement reported by Action	Significance of the output
<p>A prototypical architecture for implicitly crowdsourcing language-related data while generating vocabulary exercises was developed during the first Hackathon-like event we organized in January 2019. The code is open source and available on Gitlab. <a href="https://gitlab.com/crowdfest_task3">https://gitlab.com/crowdfest_task3</a></p> <p>The architecture has been developed with the objective of providing a generic base on top of which more implicit crowdsourcing scenarios can be built. The early publications reporting on the architecture and its first implementation were well received despite the very early stage of the work. The architecture has been tested by around 75 different users so far.</p>	High
A prototypical role-playing program which allows to	High

<p>implicitly crowdsource lexical data for non-standardized languages was created during the first Hackathon-like event we organized in January 2019. The role-playing game fosters the collaboration of the player with a person with good command of the non-standardized language (e.g. a grandparent) to answer the questions provided in the game. Despite the recent nature of the work, a demo presentation and a publication have been accepted at an international conference.</p>	
<p>A Zotero group has been created to foster the collaborative collection of references provided by members that are relevant to enetCollect. It was intensively used in relation to literature reviews during STSMs. It is currently populated with 203 references, and its use by all members will be fostered further.</p> <p>This Zotero group was also completed by crowdsourcing bibliographical information from participants of the WG1 meeting celebrated in Gothenburg at the end of 2018 as a condition for participants to join the meeting.  <a href="https://docs.google.com/spreadsheets/d/1oIbeMlyym97iC5nK4EbdFmOeRDInHx8rtubIW64w9k/edit#gid=0">https://docs.google.com/spreadsheets/d/1oIbeMlyym97iC5nK4EbdFmOeRDInHx8rtubIW64w9k/edit#gid=0</a></p>	High
<p>A dedicated channel was created on videolectures.net and video recordings of the presentations held at enetCollect meetings are being made available on the channel, including video and aligned slides.  <a href="http://videolectures.net/enetCollect/">http://videolectures.net/enetCollect/</a></p> <p>Currently 30 videos are available online, while the release of more than 40 videos of the presentations of the 2<sup>nd</sup> and 3<sup>rd</sup> Annual meetings is scheduled for the coming months.</p>	High
<p>We created a number of documents related to the COST Action workflow itself (e.g. procedures), and not to the subject of the COST Action, which we believe could be of interest for other COST-related stakeholders.</p> <p>We created a rather strict and comprehensive STSM procedure issuing STSM calls every 2 months.  <a href="http://enetcollect.eurac.edu/cost-tools-events/stsms/stsm-procedure/">http://enetcollect.eurac.edu/cost-tools-events/stsms/stsm-procedure/</a>  <a href="http://enetcollect.eurac.edu/cost-tools-events/stsms/stsm-approval/">http://enetcollect.eurac.edu/cost-tools-events/stsms/stsm-approval/</a></p> <p>We created a set of logos and templates to harmonize communications and give enetCollect an identifiable identity that members can relate to.  <a href="http://enetcollect.eurac.edu/logos_and_templates/">http://enetcollect.eurac.edu/logos_and_templates/</a></p> <p>We created a dedicated page for funding opportunities that members can pursue autonomously. The page has been quite popular (&gt;1500 visits) since it was published in Feb 2018.  <a href="http://enetcollect.eurac.edu/autonomous-funding-opportunities/">http://enetcollect.eurac.edu/autonomous-funding-opportunities/</a></p> <p>We also created guidelines about how to perform literature reviews  <a href="https://docs.google.com/document/d/128gDF-RCNIaNF5WrJMScN_cJwD3UYOVSTWurN44X5s/edit">https://docs.google.com/document/d/128gDF-RCNIaNF5WrJMScN_cJwD3UYOVSTWurN44X5s/edit</a></p>	High

<p>Two explicit crowdsourcing experiments on ranking Multi-Word Expressions (MWE) by difficulty for the purpose of creating L2 teaching materials were carried out in preparation for a WG1 hand-on workshop. Accordingly, a list of English (verbal or adverbial) multi-word expressions was associated with levels of complexity, where complexity is treated in terms of CEFR (Common European Framework of Reference) levels of proficiency (COE, 2001). The list was created through the two explicit crowdsourcing experiments where participants of a WG1 meeting were treated as a (fake) prototypical crowd of language teachers due to their high proficiency in English and their overall linguistic skills derived from their language-related R&amp;I profession.</p> <p>The aim of each crowdsourcing experiment was to collect a number of subjective ratings from crowdsourcers regarding the relative difficulty of these expressions from the perspective of a language learner. In other words, the goal was to rate expressions based on how difficult they are for a language learner to produce.</p> <p>Slides on the experiment and results:  <a href="https://spraakbanken.gu.se/sites/spraakbanken.gu.se/files/eNetCollect_WG1_GOT_JC_Experiment.pdf">https://spraakbanken.gu.se/sites/spraakbanken.gu.se/files/eNetCollect_WG1_GOT_JC_Experiment.pdf</a>  <a href="https://spraakbanken.gu.se/sites/spraakbanken.gu.se/files/Clustering%20and%20visualizing2.pdf">https://spraakbanken.gu.se/sites/spraakbanken.gu.se/files/Clustering%20and%20visualizing2.pdf</a></p> <p>Experiment:  Link to the adverbial MWE experiment: <a href="https://mnozicenje.cjvt.si/project/enetcollect_mwe_ranking_adverbs_ns/">https://mnozicenje.cjvt.si/project/enetcollect_mwe_ranking_adverbs_ns/</a>  Link to the verbal MWE experiment: <a href="https://mnozicenje.cjvt.si/project/enetcollect_mwe_ranking_verbs_ns/">https://mnozicenje.cjvt.si/project/enetcollect_mwe_ranking_verbs_ns/</a></p> <p>Link to the guidelines: <a href="https://docs.google.com/document/d/1XicAizG0nAe3gPiT47X3GWqF12SMOSyohLX7RAzV3d4/edit">https://docs.google.com/document/d/1XicAizG0nAe3gPiT47X3GWqF12SMOSyohLX7RAzV3d4/edit</a></p>	<p>Medium</p>
<p>The survey (enetCollect survey for Crowdsourcing practices, <a href="https://www.1ka.si/a/161288">https://www.1ka.si/a/161288</a>), had the MC members of the enetCollect action as a target group. The idea was to use the survey designed in Alice Millour's STSM to gather information about the crowdsourcing practices from the countries which are members of the Action. Overall information about 31 crowdsourcing resources was collected while the survey was active. The summary statistics of the survey are available in two different reports.</p> <p>Summary:  <a href="https://drive.google.com/file/d/1ia-95syaBBJzbMwkzZT5JhD9ZDh-vret/view?usp=sharing">https://drive.google.com/file/d/1ia-95syaBBJzbMwkzZT5JhD9ZDh-vret/view?usp=sharing</a>  Chart Data Analysis:  <a href="https://drive.google.com/file/d/1n2NjoZ2IR2_zu2ShVmbXxUcaxjv4kumJ/view?usp=sharing">https://drive.google.com/file/d/1n2NjoZ2IR2_zu2ShVmbXxUcaxjv4kumJ/view?usp=sharing</a></p>	<p>High</p>
<p>A survey was conducted among language teachers (N= 1129) including participants from over 30 European countries (over 35 members of the network were involved in the survey preparation and dissemination). In the survey, the following questions were addressed: (a) Are teachers familiar with the concept of crowdsourcing? (b) Are they including crowdsourcing in their teaching? (c) Are they themselves active in crowdsourcing activities? (d) If yes, what are these activities? (e) What is their attitude towards crowdsourcing in general? (f) What</p>	<p>High</p>

would motivate them to include crowdsourcing into their teaching?

While the participating teachers were generally positively inclined towards the idea of crowdsourcing, their familiarity with actual possibilities to use crowdsourcing activities in the classroom was very limited. The participants reported the need for learning about crowdsourcing as a concept as well as the need for concrete ideas about which specific crowdsourcing activities are suitable for language teaching.

Slides discussing the results

=> [https://spraakbanken.gu.se/sites/spraakbanken.gu.se/files/enetCollect-Presentation-ARHAR-Teacher\\_Survey.pdf](https://spraakbanken.gu.se/sites/spraakbanken.gu.se/files/enetCollect-Presentation-ARHAR-Teacher_Survey.pdf)

A follow up survey was designed and implemented (98 respondents) as a continuation to identify language teaching student opinions on crowdsourcing in language learning. Results have not been published yet.

The quality and quantity of the other outputs/ achievements was assessed as follows.

The enetCollect Cost Action has produced a lot of other outputs and achievements, all of them seem to be highly relevant and significant to and depend on the Action network. The first Hackathon-like event organised in January 2019 initiated a prototypical architecture for implicitly crowdsourcing language-related data while generating vocabulary exercises. Its aim is to be a generic base for future crowdsourcing scenarios. It has been tested by 75 users so far. EnetCollect has also come up with a prototypical role-playing program for implicit crowdsourcing of lexical data (non-standardized languages) which has a very good potential. Another very important output is a Zotero group that helps to collect references relevant to enetCollect. It can be used by all members of the Cost Action. It is important to note that enetCollect meetings were videorecorded and are now available on the online channel which promotes the visibility of the Cost Action. In addition, enetCollect Cost Action members produced interdisciplinary output that can be useful for other Cost Actions. They have created a compilation of the Cost Action procedures (e.g. STSM procedures, funding opportunities). This can be shared with a wider audience. The surveys and crowdsourcing experiments are also online.

## Impacts

The Action reported the following impacts (the short- to long-term scientific, technological, and / or socioeconomic changes produced by a COST Action, directly or indirectly, intended or unintended) that have resulted, or might result, from the Action.

Description of the impact	Type of impact	Timing of impact
<p>EnetCollect has managed to create a growing community of stakeholders bringing forward the new R&amp;I trend on the combination of Language Learning and Crowdsourcing. The novel nature of this community can already be attested by publications co-authored by authors who had not published together before. The growing nature of this community can be observed through the increasing number of member registering to the mailing lists and on the intranet (250 email addresses registered on the general mailing list and 160 profiles registered on the intranet) as well the growing number of people attending the Annual Meetings: 54 for the 1st in Bolzano (Italy), 68 for the 2nd in Iasi (Romania) and 105 for the 3rd in Lisbon (Portugal).</p> <p><a href="http://enetcollect.eurac.edu/about-enetcollect/statistics/">http://enetcollect.eurac.edu/about-enetcollect/statistics/</a></p> <p>This body of stakeholders is jointly pursuing several small and medium-sized research initiatives around this new R&amp;I trend, some of them already leading to publications and project proposals.</p> <p>These efforts result in a progressive sharpening of the R&amp;I trend with regards to its overall definition and objectives, as well as to its increased visibility in terms of tangible results and a referenced community of experts.</p>	<ul style="list-style-type: none"> <li>Scientific / Technological</li> </ul>	<p>Achieved</p>
<p>Validity, relevance and significance (in particular importance and timeliness) of the impact reported by the Action: The Cost Action has a high impact (Scientific/technological and societal). Based on the Action's report, it is evident that a great community of stakeholders have been created that bring forward the Research and Innovation aim of enetCollect related to language learning and crowdsourcing. New publications have come out, new members are constantly registering to the mailing list (250 emails address and 160 registered profiles on the Intranet). A large number of the participants have attended the Annual meetings. Overall: 227 members for the 3 meetings in Italy, Romania and Portugal. Cost Action members put efforts together for joint publications and research grants regarding new Research and Innovation trend. Overall, continuous work of Cost Action members leads to increased validity, relevance and significance of the Action and its impact on the society.</p>		
<p>Over the first two years, enetCollect members succeeded to secure a sizable amount of funding for enetCollect-related projects (~700 00 euros) through different small or mid-sized funding schemes. These first results allow us to feel confident regarding the capacities of enetCollect (in terms of expertise of its members and timeliness of the topic) to secure funding for follow-up projects of larger scale.</p> <p>Indeed, by building on the scientific and technological achievements reached throughout the lifetime of the Action and by exploiting established working relationships between experts of different stakeholder groups, we expect the growing number of successful cooperations to lead to a growing number of proposals for follow-up projects.</p> <p>Such new funding opportunities will allow to bring this newly established R&amp;I trend to a new level of development.</p>	<ul style="list-style-type: none"> <li>Scientific / Technological</li> </ul>	<p>Foreseen within two years of the end of the Action</p>
<p>Validity, relevance and significance (in particular importance and timeliness) of the impact reported by the Action: The scientific/ technological impact is foreseen within two years of the end of the Action. Members of enetCollect have managed to secure over 700.000 euro via enetCollect related proposals. It is a great achievement. This was possible due to the expertise of the members and the importance of the topic of the project. This will allow to have follow-up projects even after enetCollect Action comes to its end. The continuous efforts of the members of the Cost Action, growth of collaboration, networking, research, scientific and technological output will lead to growing number of</p>		

proposals and follow-up projects and will allow to develop Research and Innovation trend even further.

By performing the groundwork for integrating crowdsourcing into language learning enetCollect will enable the continuous creation and improvement of language learning material, including exercise and lesson content. In particular, applying powerful crowdsourcing approaches is expected to level up language learning material in terms of its diversification, fit to varied learner profiles and, as such, respond to the increasingly heterogeneous needs of learners resulting from the ever more varied combinations of languages taught and target groups, caused, for example, by intensified migration flows.

- Economic
- Societal

Foreseen five-to-ten years after the end of the Action

Achieving such a powerful shift towards the creation of more and more diversified language learning material will have both an economic and societal impact. On the one hand, it will strongly enhance the capacities of the stakeholders creating learning material and their opportunities for commercializing it. On the other hand, it will have the societal impact of contributing to better language learning opportunities and ultimately better language skills of all citizens and stakeholders benefiting from LL offers, a relevant variable for society.

Validity, relevance and significance (in particular importance and timeliness) of the impact reported by the Action: The economic and societal impact is foreseen within five-to-ten years after the end of the Action. The results of the enetCollect Action will have scientific, technological, economic and societal impact and great benefit to the society as integration of crowdsourcing into language learning helps to improve language learning material as well as exercises and lesson content, making them diverse and suitable for heterogeneous groups of learners of different levels of proficiency and language background. This can help to deal with an increasing migration population in Europe and world-wide. The stakeholders involved in the Cost Action will enhance their commercial benefits by creating new learning materials, while all the citizens will be offered better opportunities for language learning and improving their language skills.

By performing the groundwork for crowdsourcing on language learning enetCollect will enable the continuous creation and improvement of numerous language-related datasets. Up to now, the curation of language-related datasets has always been a cost-intensive endeavor that is often limiting the types, number of datasets and languages that can be considered and served.

- Economic
- Societal

Foreseen five-to-ten years after the end of the Action

Therefore, achieving such an impact of facilitating dataset creation by greatly lowering its costs of production will have a direct economic impact and also a potentially large catalyst effect on the R&I domains relying on such language-related datasets (e.g. NLP or corpus linguistics) for all languages learned nowadays. It will also have an indirect societal and economic impact on all the target groups benefiting from enhanced R&I performed on such language-related datasets, such as the users of these technologies or the companies commercializing them (e.g. NLP technologies).

Validity, relevance and significance (in particular importance and timeliness) of the impact reported by the Action: The economic/ societal impact is foreseen within five-to-ten years after the end of the Action. One of the benefits of the enetCollect is going to be the creation of numerous language-related datasets. This is quite an achievement as it is very time consuming and cost-intensive. EnetCollect Cost Action will allow the creation of different types of the datasets. The cost of the creation of the datasets will become lower and thus will have an economic impact on the society and will trigger the development of Research and Innovation trend (e.g. NLP corpus linguistics). The users of the datasets and technologies (e.g. NLP technologies) will have a great benefit as well.

The extent to which the Action has advanced the careers, skills and networks of researchers including ECIs (as described by the Action) is **very good**.

#### Stakeholder engagement

EnetCollect's stakeholders definition, as specified in the MoU, have evolved into more fine-grained groups in response to practical experiences in the first two years of the Action. All stakeholder groups are represented and were mostly

involved through different outreach activities (direct invitation, announcements, surveys, call for interests etc.) or learned about enetCollect through members. [Experts in language teaching] Researchers or language teachers, with knowledge of effective learning content, curricula design and needs of different learner profiles. They are well represented, especially in WG1 and WG3. [Experts in language-related data fields] Researchers with interest in and knowledge about defining, creating and annotating language-related data for research and application purposes (e.g. learner corpora, NLP). They are well represented, especially in WG2 and WG4. [Experts in LL system design and usability] Stakeholders of the Learning Management System community and Online Language-Learning community/businesses are involved in enetCollect and mostly contribute to WG3 and WG4. [Experts in crowdsourcing technologies and approaches] People from the Human Computation, crowdsourcing and GWAP communities are currently involved and mostly contribute to WG2 and WG4. [Legal, ethical and business experts] Members with knowledge about ethical/legal/business issues/opportunities related to Language Learning platforms are currently involved in WG5.

**The Rapporteur recommended that the following stakeholders should also be engaged by the Action.**

## General assessment of impacts

The Action's impacts are best described as follows:

**Multiple highly significant impacts are reasonably foreseen for the future OR one highly or moderately significant impact is already observed [Very Good]**

## Dissemination and exploitation of Action results (other than co-authored Action publications listed previously)

### Dissemination

#### Dissemination meetings funded by the Action

The following Dissemination meeting(s) funded by the Action added value for the Action:

- Dissemination meeting 2, 22-08-2018 - 25-08-2018, Finland
- Dissemination meeting 1, 07-05-2018 - 12-05-2018, Japan

#### Action website

<http://enetcollect.eurac.edu/>

The:

- openness and user-friendliness of the Action website are excellent
- content of the Action website (programmes and minutes of all events present, all outputs/deliverables accessible from website) is excellent

The Action website was an effective means of disseminating the Action.

Rapporteur's comment on the website
Very useful and efficient, with a lot of information and details included. A lot has been done during this phase of the project.

#### Other dissemination activities

The following other dissemination activities reported by the Action were effective and added value

<b>Item/activity</b>	Invited talk at the 7th DaFWEBCON - 'Webkonferenz für Deutschlehrende' (Online Conference for Teachers of German) Title: 'Sprachenlernen und Crowdsourcing – ein innovatives Projekt.' (Language Learning and Crowdsourcing, an innovative project) Presented by: Verena Lyding & Lionel Nicolas Date: March 2018
<b>Target Audience</b>	Teachers of German as second or foreign language participating on the 7th DaFWEBCON on 'Deutsch – global vernetzt – lokal aktiv!' (German - globally connected - locally active)
<b>Outcome of the activity</b>	Communication of enetCollect's ambition to teachers of German as second or foreign language. Collection of feedback and expressions of interest for being contacted for future experiments with teachers from conference participants (German teachers).
<b>Hyperlink</b>	<a href="https://dafwebkon.com/Veranstaltung/lyding-sprachenlernen-und-crowdsourcing-ein-innovatives-projekt/">https://dafwebkon.com/Veranstaltung/lyding-sprachenlernen-und-crowdsourcing-ein-innovatives-projekt/</a>

<b>Item/activity</b>	Invited talk during STSM at University of Iceland, Faculty of Languages. Title: 'Crowdsourcing in language learning and teaching: "small" languages' achievements' Presented by: Katerina Zourou Date: March 2018
<b>Target Audience</b>	Interested members of the Faculty of Languages at the University of Iceland
<b>Outcome of the activity</b>	Communication of research topics of the STSM candidate, raising awareness for the enetCollect COST Action.
<b>Hyperlink</b>	<a href="https://www.slideshare.net/Web2Learn_eu/crowdsourcing-in-language-learning-and-teaching-">https://www.slideshare.net/Web2Learn_eu/crowdsourcing-in-language-learning-and-teaching-</a>

	small-languages-achievements
<b>Item/activity</b>	Invited talk during STSM presented at the project "Industria 4.0", Corso di Letteratura italiana; Corso de Filosofia del diritto, Piacenza. Cite Title: Crowdsourc Natural Language Processing Presented by: Alain Couillault Date: May 2018
<b>Target Audience</b>	Participants of the 'Corso di Letteratura italiana' and the 'Corso de Filosofia del diritto' at the University of Piacenza
<b>Outcome of the activity</b>	Communication of project activities, raising awareness on the enetCollect COST Action
<b>Hyperlink</b>	<a href="https://www.enetcollect.net/ilias/goto.php?target=file_530_download">https://www.enetcollect.net/ilias/goto.php?target=file_530_download</a>
<b>Item/activity</b>	SPECIAL: CROWDSOURCING UND EINE REVOLUTION IM SPRACHENLERNEN at the News Section on Website of Eurac Research Authors: Verena Lyding & Lionel Nicolas Date: March 2017
<b>Target Audience</b>	Interested public, in particular stakeholders of Eurac Research in South Tyrol, Italy
<b>Outcome of the activity</b>	Increased visibility of enetCollect's objectives, raising awareness of its potential for language learning
<b>Hyperlink</b>	<a href="http://www.eurac.edu/en/news/news/pages/newsdetails.aspx?entryid=122107">http://www.eurac.edu/en/news/news/pages/newsdetails.aspx?entryid=122107</a>
<b>Item/activity</b>	Two short radio transmissions at 'Morgenmagazin „Thema Wissen“' (Morning show on 'Knowledge/Science'), RAI Südtirol Topics: 'Crowdsourcing and Language Learning' and 'Computer-based Language Learning' Invited interview partner: Verena Lyding Date: June 2017
<b>Target Audience</b>	General public in South Tyrol, Italy, interested in transmissions on knowledge and science
<b>Outcome of the activity</b>	Communication of the project objectives, raised awareness on the crowdsourcing potential for language learning.
<b>Hyperlink</b>	<a href="http://www.raibz.rai.it/index.php">http://www.raibz.rai.it/index.php</a>
<b>Item/activity</b>	Presentation at mini-workshop collocations and sentence selection on the creation of language learning materials and automatic sentence selection for learner reference resources at Språkbanken, University of Gothenburg during an STSM. Title: "Semi-automated lexicography in Slovenia: collocations dictionary and beyond" Presented by: Iztok Kosem Date: May 2017
<b>Target Audience</b>	Researchers from the research group at Språkbanken, University of Gothenburg, as well as external research fellows.
<b>Outcome of the activity</b>	Communication and discussion of automatic data extraction methods used in Slovene lexicography and language learning; and presentation of plans for implementing crowdsourcing techniques into the workflow
<b>Hyperlink</b>	<a href="https://sweclarin.se/swe/mini-workshop-collocations-and-sentence-selection">https://sweclarin.se/swe/mini-workshop-collocations-and-sentence-selection</a>
<b>Item/activity</b>	Invited presentation at the research group for Second Language Studies at the University of Iceland during an STSM. Title: "The scope of language intuition: some insights from crowdsourcing" Presented by: Karën Fort Date: November 2017
<b>Target Audience</b>	Member of the research group for Second Language Studies at the University of Iceland, <a href="http://rim.hi.is">http://rim.hi.is</a>
<b>Outcome of the activity</b>	Communication of relevant crowdsourcing work of the STSM applicant to researchers at the host institution.
<b>Hyperlink</b>	<a href="https://enetcollect.net/ilias/goto.php?target=file_379_download">https://enetcollect.net/ilias/goto.php?target=file_379_download</a>

<b>Item/activity</b>	Presentation for researchers from the tagged Icelandic corpus team ( <a href="http://www.malfong.is/index.php?lang=en&amp;pg=mim">http://www.malfong.is/index.php?lang=en&amp;pg=mim</a> ) Topic: Demonstration of ZombiLingo (dependency syntax), Bisame (POS tagging) and Rigor Mortis (multi-word expressions) Presented by: Karën Fort Date: November 2017
<b>Target Audience</b>	Researchers from the tagged Icelandic corpus team ( <a href="http://www.malfong.is/index.php?lang=en&amp;pg=mim">http://www.malfong.is/index.php?lang=en&amp;pg=mim</a> ) interested in crowdsourcing examples.
<b>Outcome of the activity</b>	Communication about three existing crowdsourcing tools (GWAPs) for language resources. Participants to the presentation showed some interest and could be willing to collaborate on a Game with a purpose for Icelandic multi-word expressions or to adapt ZombiLingo.
<b>Hyperlink</b>	<a href="https://enetcollect.net/ilias/goto.php?target=file_379_download">https://enetcollect.net/ilias/goto.php?target=file_379_download</a>

<b>Item/activity</b>	Presentation at the Economics Department and Business School of the University of Iceland. Title: "Contractual governance and performance Presented by: Nina Gorovaia Date: January 2018
<b>Target Audience</b>	Researchers and members of the Economics Department and Business School of the University of Iceland.
<b>Outcome of the activity</b>	Communication of research of the STSM grantee for networking. Very useful contacts have been made that might result in future collaboration and consulting opportunities for the public sector.
<b>Hyperlink</b>	<a href="https://www.enetcollect.net/ilias/goto.php?target=file_416_download">https://www.enetcollect.net/ilias/goto.php?target=file_416_download</a>

<b>Item/activity</b>	Presentation to the members of Språkbanken, the Swedish Language Bank, University of Gothenburg (SB research meeting series) Topic: Previous work on exercise generation Presented by: Itziar Aldabe Date: March 2018
<b>Target Audience</b>	Researchers and members of Språkbanken, University of Gothenburg, interested in exercise generation.
<b>Outcome of the activity</b>	Communication of research work by the STSM grantee relevant for WG1.
<b>Hyperlink</b>	<a href="https://www.enetcollect.net/ilias/goto.php?target=file_518_download">https://www.enetcollect.net/ilias/goto.php?target=file_518_download</a>

<b>Item/activity</b>	Guest talk at the University of Ljubljana and Trojina Institute Title: "Towards a research infrastructure for Second Language Acquisition and teaching: case of L2 Swedish" Presented by: Elena Volodina Date: June 2018
<b>Target Audience</b>	A good portion of audience represented the highly needed group of L2 researchers, teachers and assessors.
<b>Outcome of the activity</b>	Communication of research activities of the STSM grantee to a relevant target group; networking and creation of opportunities for involvement into WG1.
<b>Hyperlink</b>	<a href="https://www.enetcollect.net/ilias/goto.php?target=file_544_download">https://www.enetcollect.net/ilias/goto.php?target=file_544_download</a>

<b>Item/activity</b>	Presentation at meeting with DALI team (Disagreements and Language Interpretation) at Queen Mary University at London (QMUL academic leader -Massimo Poesio). Topic: Presentation of enetCollect, WG1-5, Smart User Model and Linguistics Apps Development for Second Language Learning Presented by: Rina Zviel-Girshin Date: September 2018
<b>Target Audience</b>	Members of the DALI team (Disagreements and Language Interpretation) at Queen Mary University at London.
<b>Outcome of the activity</b>	Communication of the enetCollect Action and ongoing research of the STSM grantee.
<b>Hyperlink</b>	<a href="https://www.enetcollect.net/ilias/goto.php?target=file_658_download">https://www.enetcollect.net/ilias/goto.php?target=file_658_download</a>

<b>Item/activity</b>	Post-meeting proceedings for the WG3/WG5 meetings celebrated in Leiden at the end of
----------------------	--

	2019.
<b>Target Audience</b>	The target audience is composed of all stakeholder groups involved in enetCollect, especially the ones related to WG3 and WG5.
<b>Outcome of the activity</b>	The proceedings should be published during the summer 2019 and are thus not available yet. Nonetheless, a call for paper was launched, 13 submissions were made and reviewed by two peers. 12 papers were selected for publication.
<b>Hyperlink</b>	<a href="https://groups.google.com/d/msg/enetcollect-all/tJH4p74OC6w/hlrE95aTBQAJ">https://groups.google.com/d/msg/enetcollect-all/tJH4p74OC6w/hlrE95aTBQAJ</a>

<b>Item/activity</b>	Presentation at IXA seminar series, University of the Basque Country. Topic: Towards a lightweight solution for less-resource languages Presented by: Alice Millour Date: February 2018
<b>Target Audience</b>	Researchers and members of the IXA NLP research group, University of the Basque Country.
<b>Outcome of the activity</b>	Communication of research work by the STSM grantee relevant for WG2.
<b>Hyperlink</b>	<a href="http://www.enetcollect.net/ilias/goto.php?target=file_802_download&amp;client_id=enetcollect">http://www.enetcollect.net/ilias/goto.php?target=file_802_download&amp;client_id=enetcollect</a>

<b>Item/activity</b>	Professional video recording of presentations held at enetCollect meetings and creation of a channel on videolectures.net to publish and archive the videos together with slides.
<b>Target Audience</b>	Action members as well as any interested stakeholders of the Action.
<b>Outcome of the activity</b>	Dissemination of the presentations held at enetCollect meetings to a wider audience. Formation of interested stakeholders on enetCollect related topics. Currently, 30 videos are available online and ca. 40 more videos are in processing.
<b>Hyperlink</b>	<a href="http://videolectures.net/enetCollect/">http://videolectures.net/enetCollect/</a>

<b>Item/activity</b>	Creation of a Zotero group (reference management software) for the collaborative collection of references relevant to enetCollect's domains of interest.
<b>Target Audience</b>	EnetCollect members as well as any stakeholders of the Action; researchers and students of the related research domains.
<b>Outcome of the activity</b>	Centralized and collaboratively curated collection of relevant related work on the R&I trend on Language Learning and Crowdsourcing.
<b>Hyperlink</b>	<a href="https://www.zotero.org/groups/2174884/_enetcollect_bibliography_">https://www.zotero.org/groups/2174884/_enetcollect_bibliography_</a>

<b>Item/activity</b>	Several announcements about the enetCollect Action as well as multiple transfer-oriented presentations and articles were published by members on their local/institutional webpages etc.
<b>Target Audience</b>	Local research communities of members, interested public, any visitors of members' webpages.
<b>Outcome of the activity</b>	Information about enetCollect, its objectives, community and preliminary outputs is communicated to a wider (also non-scientific) audience.
<b>Hyperlink</b>	<a href="http://enetcollect.eurac.edu/results/general_dissemination/">http://enetcollect.eurac.edu/results/general_dissemination/</a>

<b>Item/activity</b>	Presentation at the 4th Conference for Blended Learning, Zayed University, Dubai. Title: European Network for Combining Language Learning and Crowdsourcing Techniques – a New Perspective in Blended Learning. Presented by: Elżbieta Gajek Date: April 25th-27th, 2019
<b>Target Audience</b>	Stakeholders of the International Association for Blended Learning in Partnership with the Center for Educational Innovation, Zayed University, Dubai.
<b>Outcome of the activity</b>	Promoting the enetCollect Action to professionals and practitioners related related to blended

	learning.
<b>Hyperlink</b>	<a href="https://iabl.wildapricot.org/resources/Documents/WCBL_Proceedings.pdf">https://iabl.wildapricot.org/resources/Documents/WCBL_Proceedings.pdf</a>

## Exploitation activities

The following activities to ensure exploitation (use, in particular in a commercial context) of the Action's achievements reported by the Action were effective and added value

<b>Item/activity</b>	WG 2 together with members of the Core Group organized an Hackathon-like initiative in January 2019 to kick-start the implementation or theoretical elaboration of approaches/experiments/applications of relevance for the different WGs. The Core Group intends to repeat this successful event in 2020.
<b>Target Audience</b>	Any members of enetCollect interested in making a concrete and practical contribution to a subject of shared interest among various members, possibly from different WGs.
<b>Outcome of the activity</b>	During two days, more than 35 members worked in subgroups targeting 6 tasks related to different WGs. A number of tangible outputs resulted from such efforts, including various prototypes and publications.

<b>Item/activity</b>	During the Hackathon-like event organized in January 2019, one of the 6 tasks on which the members worked on was related to "Business models in language learning platforms" and aimed at exploring the business opportunities that enetCollect could create over time by first studying as a reference point the business models of existing platform.
<b>Target Audience</b>	Any members of enetCollect with a business-oriented mindset and interested in creating commercial solutions from the developments made within enetCollect.
<b>Outcome of the activity</b>	The preliminary results achieved were presented during the 3rd Annual Meeting in Lisbon and an STSM proposal will be submitted to extend this work further.

## Assessment of Action dissemination and exploitation activities

The effectiveness of the Action's dissemination and exploitation approach (other than co-authored publications) is assessed as follows

Dissemination and Exploitation Approach of the Cost Action is very effective. The coordination groups of the enetCollect Cost Action offer support and opportunities for growth of relevant initiatives without any specific limitations. EnetCollect uses various means of communication such as websites, mailing lists, newsletters, social media. This helps to transmit relevant initiatives and achievements to large audiences (in and outside Cost Action). Action-internal dissemination include presentation of enetCollect meetings. Press releases are used for Action-external dissemination. Regarding the exploitation of the enetCollect's results, it is expected that a large group of professional language service providers (language learning platforms, translation services) as well as business partners, associations, non-profit organisations, academics and schools will find the results useful. EnetCollect also promotes open access data. Overall, it shows that the Dissemination and Exploitation Approach of the Cost Action will be efficient and have long-time impact.

Assessment of Action dissemination and exploitation activities

There were many highly effective Action activities focusing on dissemination of Action results [Excellent]  
There were many highly effective Action activities focusing on exploitation of Action results [Excellent]

## Other matters

### Difficulties in implementing the Action

The Action Rapporteur made the following observations regarding difficulties in implementing the Action:

There are some difficulties in implementing the Action, but these can be overcome in the next phases of the Cost Action.

Most of the concerns are related to CG, the Chairs and the members. Some CG members were underperforming at the beginning of the Action due to some misunderstanding, confusion and openness of the Cost Action approach, but the Chair tried to improve the situation via specific sub-initiatives on a technical, procedural and managerial level: STSMs, inviting relevant experts.

This at the same time created a bottleneck situation due to the limitation of time and high workload on the side of the Chair. The Chair was investing more time and work but also tried to transfer and distribute sub-tasks to single members in order to make the procedures more efficient.

It was also decided to provide Cost Action funding only to active MC members who contribute to the research output and collaboration meetings and not the passive ones.

Overall, these initiatives can help to overcome difficulties in implementing the Action.

### Emerging topics / developments in the field of the Action

The Action reported the following emerging topics / developments in the field of the Action.

- By emerging topics that could be addressed by future COST activities, we understand content-transversal topics that are of interest for several COST Actions.

Related to enetCollect, the following topics could be identified which could be addressed by future COST activities:

- ethics and legal issues related to gathering research data (e.g. to discuss GDPR-related questions),
- methodological issues related to citizen science approaches (e.g. to discuss citizens' profiles and motivations for participating),
- big data curation and storage (e.g. to discuss best-practice in terms of metadata, long term sustainability etc.).

The Action Rapporteur made the following comment on the emerging topics / developments in the field reported by the Action.

The emerging topics and potential future developments identified by the Action are important and sustainable.

The future Cost Activities are going to include content-transversal topics that are of great importance to several Cost Actions, such as ethics and legal issues with respect to how to gather research data in a proper way; methodological issues regarding citizen science approaches; big data curation and storage.

These topics are interdisciplinary and are related to data collection, analysis and use as well as ethical issues and impact on the society.

## Action Rapporteur

This Second Progress Report was submitted on 2019-07-09 by:

Dr Sviatlana Karpava

University of Central Lancashire, Cyprus

Cyprus

## Annex 1: List of publications

The Action reported 9 publications on the topic of the Action, co-authored by at least two Action participants from two countries participating in the Action.

### Co-authored Action publications - peer-reviewed

Title	Parafraseamento Automático de Registo Informal em Registo Formal na Língua Portuguesa
Authors	Anabela Marques Barreiro; Ida Rebelo-Arnold; Jorge Baptista; Cristina Mota; Isabel Garcez
DOI	<a href="https://doi.org/10.21814/Im.10.2.282">doi:10.21814/Im.10.2.282</a>
Type	Journal article
Published in	Linguamática
Published by	University of Minho
ISSN	<a href="https://issn.org/1647-0818">1647-0818</a>
Link	<a href="http://linguamatica.com/index.php/linguamatica/article/download/282/447">http://linguamatica.com/index.php/linguamatica/article/download/282/447</a>
Title	enetCollect: A New European Network for combining Language Learning with Crowdsourcing Techniques
Authors	Rodrigo Agerri; Montse Maritxalar; Verena Lyding; Lionel Nicolas
DOI	<a href="https://doi.org/10.26342/2018-61-25">doi:10.26342/2018-61-25</a>
Type	Journal article
Published in	Procesamiento del Lenguaje Natural
Published by	Sociedad Española para el Procesamiento del Lenguaje Natural
ISSN	<a href="https://issn.org/1989-7553">1989-7553</a>
Title	Introducing the European NETwork for COmbining Language LEarning and Crowdsourcing Techniques (enetCollect)
Authors	Verena Lyding; Lionel Nicolas; Branislav Bédi; Karën Fort
DOI	<a href="https://doi.org/10.14705/rpnet.2018.26.833">doi:10.14705/rpnet.2018.26.833</a>
Type	Chapter
Published in	Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018
Published by	Research-publishing.net
Link	<a href="https://research-publishing.net/publication/chapters/978-2-490057-22-1/833.pdf">https://research-publishing.net/publication/chapters/978-2-490057-22-1/833.pdf</a>

- Nicolas, Lionel, Verena Lyding, Luisa Bentivogli, Federico Sangati, Johanna Monti, Irene Russo, Roberto Gretter, and Daniele Falavigna. "EnetCollect in Italy." In *Proceedings of the Fifth Italian Conference on Computational Linguistics*. Torino, 2018.
- Zvieli-Girshin, Rina, Branislav Bedi, and Nathan Rosenberg. "EnetCollect User Model for Intelligent

- Gamification Crowdsourcing Language Learning IS.” In *Proceedings of the 12th Israel Association for Information Systems (ILAIS) Conference*. Ashdod, 2018.
- Zviel-Girshin, Rina, Nathan Rosenberg, and Christos Rodosthenous. “SUSTAIN-ALL Language Learning.” In *Proceedings of the 12th Israel Association for Information Systems (ILAIS) Conference*. Ashdod, 2018.
  - Sangati, Federico, Ekaterina Abramova, and Johanna Monti. “DialettiBot: A Telegram Bot for Crowdsourcing Recordings of Italian Dialects.” In *Proceedings of the Fifth Italian Conference on Computational Linguistics*. Torino, 2018.
  - Alice Millour, Marianne Grace Araneta, Ivana Lazić Konjik, Annalisa Raffone, Yann-Alan Pilatte and Karën Fort, "*Katana and Grand Guru: a Game of the Lost Words (DEMO)*" in *Proceedings of the ninth Language & Technology Conference*, Poznan, Poland, 2019.
  - *Designing a Prototype Architecture for Crowdsourcing Language Resources*, Christos Rodosthenous and Verena Lyding and Alexander König and Jolita Horbacauskiene and Anisia Katinskaia and Umair ul Hassan and Nicos Isaak and Federico Sangati and Lionel Nicolas, in *Proceedings of the 2nd Language, Data and Knowledge (LDK)*, 2019.