Sign Language Translation Mobile Application and Open Communications Framework

Deliverable 7.6: Second Impact Review

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101017255

Project Information
Project Number: 101017255
Project Title: SignON: Sign Language Translation Mobile Application and Open Communications Framework
Funding Scheme: H2020 ICT-57-2020
Project Start Date: January 1st 2021

Deliverable Information
Title: Second Impact Review
Work Package: WP7: Coordination and Management
Lead beneficiary: DCU
Due Date: 31/12/2022
Revision Number: V1.0
Authors: Aoife Brady, Dimitar Shterionov, contributions from the whole consortium
Dissemination Level: Public
Deliverable Type: Report

Overview: This report provides an overview of the impact of the SignON project to date. It gives an outline of the impact that was described in the Grant Agreement and how this impact has been seen at the end of the first year of the project.

Revision History

Version #	Implemented by	Revision Date	Description of changes
V0.1	Aoife Brady	17/11/2022	Initial Draft
V0.2	Dimitar Shterionov	18/12/2022	Revised Draft
V1.0	Dimitar Shterionov, Aoife Brady	20/12/2022	Final Version

The SignON project has received funding from the European Union's Horizon 2020 Programme under Grant Agreement No. 101017255. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the SignON project or the European Commission. The European Commission is not liable for any use that may be made of the information contained therein.

Approval Procedure

Version #	Deliverable Name	Approved by	Institution	Approval Date
V0.1	D7.6	Shaun O'Boyle	DCU	13/12/2022
V0.1	D7.6	Vincent Vandeghinste	INT	06/12/2022
V0.1	D7.6	Gorka Labaka	UPV/EHU	09/12/2022
V0.1	D7.6	John O'Flaherty	MAC	22/11/2022
V0.1	D7.6	Víctor Ubieto	UPF	09/12/2022
V0.1	D7.6	Lorraine Leeson	TCD	09/12/2022
V0.1	D7.6	Jorn Rijckaert	VGTC	13/12/2022
V0.1	D7.6	Ellen Rushe	NUID UCD	09/12/2022
V0.1	D7.6	Henk van den Heuvel, Louis ten Bosch	RU	6/12/2022
V0.1	D7.6	Catia Cucchiarini	TaalUnie (NTU)	19/12/2022
V0.1	D7.6	Myriam Vermeerbergen	KU Leuven	09/12/2022
V0.2	D7.6	Davy Van Landuyt	EUD	19/12/2022
V0.1 V0.2	D7.6	Dimitar Shterionov Dimitar Shterionov	TiU	05/12/2022 18/12/2022

Acronyms

The following table provides definitions for acronyms and terms relevant to this document.

Acronym	Definition
API	Application Programming Interface
ASR	Automatic Speech Recognition
CLARIN	Common Language Resources and Technology Infrastructure
DHH	Deaf and Hard of Hearing
DoA	Description of Action
GA	Grant Agreement
InterL	Interlingua
IP	Intellectual Property
mBART	Multilingual Bidirectional and Auto-Regressive Transformer
MT	Machine Translation
NLU	Natural Language Understanding
SL	Sign Language/s
SLR	Sign Language Recognition
SLTAT	Sign Language Translation and Avatar Technology Workshop
SOTA	State-Of-The-Art
SUS Rating	System Usability Scale Rating
UI	User Interface
WP	Work Package

Table of Contents

1. Executive Summary

Our society is phonocentric - that is, built around the assumption that sound and speech are more primary than other forms of language use (writing; and languages expressed in a visual-gestural modality, sign languages). As a result of this auditory bias, people who are deaf or hard of hearing (DHH) are often excluded from accessing information in a society which is frequently communicated in an auditory modality (audio announcements, videos with no captions, etc.). Additionally, there is the presumption that captions solve the modality communication gap - but we must remember that the spoken and written languages of Europe are linguistically distinct from the signed languages of Europe. Thus, many signers experience the written languages of their home country as second languages, and literacy in these languages varies. We can also say that hearing people face challenges in trying to communicate with people in DHH communities too, but while this may be an equivalent communicative challenge in the moment of communication, the challenge is not the same in terms of cumulative effort expected of deaf people to bridge communication gaps with hearing people. This burden is asymmetrical and impacts DHH people significantly more in their everyday lives, over the course of their lives. To try to address this challenge, the EU-funded SignON project aims to develop a mobile application that will translate between different European signed and spoken languages.

The application, which has the form of lightweight software running on a modern mobile device with a camera, will interact with a cloud-based distributed framework dedicated to these computationally heavy tasks. The application and the framework are designed through a co-creation approach where DHH and hearing users work together with SignON researchers and engineers to build a solution that suits all user types. Furthermore, it is built for easy adaptability to other languages (signed and spoken) and modalities other than text, i.e. audio and visual. Ultimately, the application will promote equitable exchange of information among all European citizens.

This project runs over three years from January 2021 to December 2023, and at the time of writing this report, we are almost 24 months into the execution of the project. In this review, we present the impact of the project to date.

2. Introduction

One in a thousand people in the EU (approximately half a million DHH persons) communicates in one of 31 national or regional sign languages (SL) as their first or preferred language. The SignON service aims to reduce the communication gap between deaf communities, people who are hard of hearing (including people with cochlear implants), hearing people (including family members and friends of deaf people), industry partners and policy maker user groups and promote inclusion of deaf communities into society by providing an open source and scalable solution for real-time translation between SL, text and speech in the languages that will be addressed in the project (Flemish Sign Language (VGT), Sign Language of the Netherlands (NGT), Irish Sign Language (ISL), British Sign Language (BSL) and Spanish Sign Language (LSE), as well as English, Irish, Dutch & Spanish spoken languages). With the expected wide take-up of the technology, SignON will make a significant contribution to its users and the community. SignON contributes to the repertoire of communication means and tools between deaf, hard of hearing and hearing people.

Currently, the automatic recognition of naturally performed SL is a challenging issue due to its modality. Both manual and non-manual features are used in parallel to construct a signed utterance and their identification is crucial for real-time recognition. To build efficient data-driven Sign Language Recognition (SLR) models, large datasets are required and the lack of such resources results in poor applicability of currently available systems.

SignON will also impact academic research, industry innovation and business progression through its SOTA advancements in real-time SL linguistics, SLR, Natural Language Understanding (NLU), Automatic Speech Recognition (ASR) (including atypical speech), Machine Translation (MT) and 3D avatar technology, and their underpinning fundamentals. Our platform and its constituent tools, will offer an alternative vision for the governance and values of SL translation and will be publicly available to end-users in the form of a mobile app for real-time translation. SignON is a fundamental step towards a new vision for SLR and translation between sign and spoken languages. It employs a user-centric approach to draw use-cases, to validate quality, to ensure acceptance based on continuous communication with the stakeholders. As such SignON aims to become a stepping stone to foster communication, information exchange, business creation, learning and knowledge acquisition with an inclusive and open society.

3. Expected Impacts as Defined in the Grant Agreement

The expected impacts outlined in the Grant Agreement (GA) address the expectations set by the programme call and were embedded in the original proposal (which later became the GA). Additional impacts were defined, including impacts specific to each consortium partner. These expected impacts are described below and the impact of each to the end of the second year of the project is outlined. Some of the wording of these is revised to make it more clear as they are taken outside of the context of the Grant Agreement. This is explicitly noted as a footnote along with the revised impact.

3.1 Expected impacts of the Project

SignON results will achieve impacts at multiple levels for the user communities, business, industry and research stakeholders as well as the consortium partners, which we will discuss in sections 3.1.1 (Expected Impacts as per the Call), 3.1.2 (Additional Impacts) and 3.1.3 (Impacts for the Consortium).

3.1.1 Expected Impacts as per the Call

Improve multilingual speech processing on mobile devices:

SignON will improve multilingual speech processing by:

1. Personalisable ASR for typical and atypical speech for English, Dutch, Spanish and Irish, accessible via the SignON application;
2. Enhancement and processing of the ASR output through language-specific NLU pipelines for improved message understanding and translation;
3. Exploiting a multilingual, multi-modal intermediate representation to capture the meaning of a message regardless of the source language and modality;.
4. ASR tuned to DHH speakers (atypical speech), and to our knowledge SignON is the first to provide this. Recent evaluations of ASR for the DHH community could only report about performances for standard ASR engines[1]. Thus SignON makes a significant step forward in inclusiveness. The same holds for the ASR and NLU facility for an under-resourced language like Irish. The ASR models SignON will develop will be optimised for mobile devices and will be accessible online (communicating with a remote server hosting ASR models) or offline (models residing on the mobile device).

[1] Glasser, A. (2019) Automatic speech recognition services: Deaf and Hard-of-Hearing usability. C HI'19.

Impact Progress Indicators from the Grant Agreement:

1. Improved speech to text recognition assessed by automatic metrics over all SignON supported spoken languages;
2. Increased preference in using speech to text by the users of SignON, for all SignON supported languages - automatically assessed via the SignON app;[2]
3. User satisfaction of translation based on speech input via voluntary user feedback based on a Likert 5-point scale.
4. User satisfaction of mobile services of at least 68.0[3] System Usability Score (SUS).

Additional Progress Indicators

As a result of the MidTerm Review, we will continue to actively progress on these indicators, and in addition we will focus on development of the SignON SLR (Sign Language Recognition) and SLS (Sign Language Synthesis, i.e. Avatar) functionality. We will also focus on developing demonstrations of the latest versions of the SignON App and facilitating potential users' structured feedback and suggestions for its UI (as in D1.9) at co-creation events to achieve the following key progress indicators:

1. A new Milestone (MS11) demonstrating the SignON App with SLR and SLS services at TRL 5 (technology validated in relevant environment[4]) by March 2023.
2. An enhanced final Milestone (MS12) demonstrating the SignON App with operational SLR and SLR services at TRL7 (system prototype demonstration in operational environment) by the end of the project in December 2023.

Progress Indicators to Date:

[2] This impact indicator is a revised version of the one mentioned in the grant agreement "Increased preference in using speech input to text reaching (at the end of the project) over all SignON supported languages - automatically assessed via the SignON app" in order to make it more clear as it is taken outside of the context of the grant agreement.

[3] 68.0 is a widely accepted standard; https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

[4] h2020-wp1415-annex-g-trl_en.pdf (europa.eu)

We have developed and implemented the first ASR models for Dutch[5], English and Spanish. For Dutch, we have realised two variants, one for Northern Dutch (as spoken in the Netherlands) and one for Southern Dutch (Flemish, as spoken in Belgium). The main reason for this is the considerable difference in pronunciation which requires different acoustic models. In our tests the developed ASR components achieve word error rates (WERs) of around 15%. For Irish, being a seriously under-resourced language, a transfer-learning approach using wav2vec2[6] is followed, in which we borrow the acoustic components from other languages, whereas we compile the language model from existing texts. Following the same approach, we have developed a first version for Dutch atypical speech of deaf speakers and speakers with cochlear implants, which will be tested in the next period.

Prior to releasing these models and their integration into the App, the App used Google speech to text. In D1.9 the user community rated the SignON App V1.0 speech to text (using Google Translate) to be easy to use and their overall SUS rating for the SignON Mobile App is just below the SUS threshold of acceptability of 68. Both SL users (63) and other users (62.5) rated it very similarly, but both under 68. This is to be expected as only the speech and text translations are functional at present and there are currently no SL recognition functions in place. However, these SUS scores indicate that the SignON App V1.0 prototype has started on the right track with regard to what users want and need.

Improve sign language recognition on mobile devices:

SignON will deliver a mobile application with support for SLR based on single-camera video. While other technologies for SLR use excessive hardware (e.g. multiple cameras, wearables, etc.), SignON will fuse together state-of-the-art AI technologies - for face, hand and posture estimation - in an SLR component with improved performance. The SLR component will be based in the cloud, making the application available on any modern mobile device with a camera, without the need for specialised hardware or customised phones. Its distributed architecture will allow for parallel processing of computationally expensive tasks on high-end servers, leading to reduced latency alongside the high recognition accuracy. As SignON will support multiple SLs (see below), and because it will not require expensive special equipment it will be a highly inclusive solution for SLR on mobile devices.

[5] A punctuation / segmentation model for Dutch and also a multilingual model have been developed as well. They are, available on HuggingFace: https://huggingface.co/oliverguhr/fullstop-dutch-sonar-punctuation-prediction and https://huggingface.co/oliverguhr/fullstop-punctuation-multilingual-sonar-base and described in Vandeghinste & Guhr (2022). FullStop: Punctuation and Segmentation Prediction for Dutch with Transformers.

[6] Baevski, A., Zhou, H., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. In Proceedings of NeurIPS 2020

Impact Progress Indicators from the Grant Agreement:

1. Improved SLR according to automatic metrics, e.g. accuracy, comparable to academic SOTA results under similar conditions;
2. Increased preference in using SL recognition to text input or speech recognition (for signers who can speak) for all supported SLs - automatically assessed via the SignON app;
3. User satisfaction of translation based on SL input via voluntary user feedback based on a Likert 5-point scale.
4. User satisfaction of mobile services of at least 68.0 System Usability Score (SUS).

Progress Indicators to Date:

Following SOTA principles and practices, we have set up an initial SLR component. This component can extract and post-process human poses using state-of-the-art techniques for body, face and hand pose estimation. These kinds of data have been evaluated for sign language translation (SLT) (on a dataset that is not part of the project, because dataset collection for SignON was still ongoing at the time), with promising initial results. Currently, at the end of year 2, we have trained SLR models for NGT, VGT, BSL and ISL. We are still working on models for LSE. LSE, like BSL and ISL is one of the languages with very limited resources. In fact, there are a lot of resources for BSL, however few are openly available. Unlike BSL and ISL, utilising the LSE data is harder and needs more work as the available data is not completely labelled. We plan to work on LSE for the 3rd year of the project, where we have scheduled the development of multilingual SLR. We have started from the open-source model MediaPipe and we have built SL-specific mechanisms to deal with errors and improve the SLR quality. The component has been integrated in the SignON framework, facilitating the translation from SL to text (D4.9). In the coming year we first plan to conduct experiments to quantify the quality of sign-to-text translation pipeline and, second, additional work will be conducted to improve the SLR component with respect to this task for all supported languages. In the coming year, this will be updated incorporating findings from experiments and folding in results from linguistic analysis of the sign languages in our projects.

Deploy solutions allowing wide take up by people who are deaf or hard of hearing:

Many past and current technologies and services for DHH people have been unsuccessful due to failure to meet users' requirements and/or expectations or due to the artificial setting of design and development, as well as due to the lack of available data sets. Through its co-creation workflow, SignON will positively impact user communities by responding to their self-selected needs for real-life

communication. The users will drive this project and as a result, we will see the weight and importance of their decisions imprinted in the SignON service. Furthermore, our user-centric practices will set an example of how applications and services for accessibility and inclusiveness should be developed. In such a way, SignON will establish a standard for the development of user-centric technology through co-creation methodologies.

Impact Progress Indicators from the Grant Agreement:

1. Constant increase in the number of active users (of intermediate and final releases) and users who have expressed interest in the final release based on demonstrations of and/or own experience with the intermediate releases and decrease in potential users with negative attitudes towards the proposed technology (measured through co-creation events, figures expressed in both absolute and relative values);
2. Constant increase in the number and demographic spread of social media followers from the DHH cohorts and an increase in social media activity.

Progress Indicators to Date:

The SignON project, through its communication and dissemination as well as thorough co-creation activities has achieved significant social impact which we measure by the number of events held or invited to, number of users and continuous user engagement. In terms of presentations and participations in various forums (including academic conferences and workshops), in the second year of our project, SignON work was presented at NITS2022, EAMT2022[7], LREC2022[8], CLIN32[9], NLDB[10], SLTAT2022[11] and the 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources[12]. It was also the subject of a presentation for the members of the Flemish and Dutch parliaments, and at the General Assembly of the EUD where members from each of the of the European National Deaf Associations were present. Through our social media channels, we have gathered a significant number of users. At the end of the second year (November 20, 2022) we had c. 3,500 unique visitors and 10,600 pageviews to the project website, 564 Facebook followers, 419 Twitter followers, 113 LinkedIn followers and 155 Instagram followers. We published an item about the

[7] https://eamt2022.com
[8] https://lrec2022.lrec-conf.org/en/
[9] https://clin2022.uvt.nl/
[10] https://nldb2022.prhlt.upv.es
[11] http://sltat.cs.depaul.edu/sltat_2022.htm
[12] https://www.sign-lang.uni-hamburg.de/lrec2022/programme.html

availability of two VGT corpora in INT's newsletter (4300 subscribers) of November 2022,[13] and an associated page explaining the BeCoS corpus,[14] both targeting the general public. SignON, in cooperation with its sister project EASIER is making sure that sign languages are included in the European Language Equality initiative (ELE),[15] by writing ELE deliverable 1.40 *Report on Europe's Sign Languages*, which is targeting decision makers at the European and national/regional levels, language communities, journalists, etc. The ELE related initiative was mentioned and welcomed by participants in the STOA workshop *Towards full digital language equality in a multilingual European Union*, held at the European Parliament on 8th of November 2022.

Addressing under-resourced languages:

SLs are under-resourced and are under-analysed relative to spoken languages. There is a need to develop resources in SLs and to describe the languages of interest more fully as we go. The SignON project will strengthen the position of the supported SLs by contributing to the documentation of sign languages and by developing advanced technology that will allow different forms of translation between languages and modalities. The SignON framework will include automated mechanisms and methodologies for expanding, improving and verifying the service. It will facilitate the inclusion of new languages via the SignON technology for learning from user content which is intended for low-resource scenarios or, thanks to SignON's modular design, via new and upgraded tools and models. Through SignON's learning capability, new resources can be collected and existing ones verified. The interactive 3D-puppeteering we will develop within this project will allow, through co-creation, to indefinitely expand the database of signs, validate and increase the quality of the automatic generation of Virtual Signers. In combination with the resources that can be collected through the SLR and ASR, SignON can create new parallel corpora. Extending the corpora pool will reduce a significant gap for data-driven methods for recognition of sign and speech. As such it will impact the advancement in this trend of research. In this way the SignON project will not only impact the users of the SignON service, but also the research and development communities and enable them to advance the SOTA in ASR, SLR, avatar technologies, SL translation and synthesis.

Impact Progress Indicators from the Grant Agreement:

1. Positive user assessment of the QoS for low-resource languages such as Irish;

[13] https://mailchi.mp/ivdnt/algemenenieuwsbrief_november2022
[14] https://ivdnt.org/actueel/taalmateriaal-uitgelicht/becos-corpus-covid-19/
[15] https://european-language-equality.eu/

2. Validation of the machine learning technology of SignON by the academic (citation number of SignON publications) and the user community;
3. Deploying new or extending existing corpora of speech-to-text, SL-to-text and SL-to-speech corpora as well as database of SL (monolingual data) for Virtual Signers research and development.

Progress Indicators to Date:

The SignON Mobile App V1.0 prototype's speech to text offers both Google Translate and SignON ASR and is available on both Android and Apple phones. The distributed approach we take allows for updating, improving or replacing certain tools and models (as stated above, for example, one can choose to use either Google Translate or SignON ASR). Furthermore, the App can easily switch to the models we develop specifically for the needs of SignON. These models can be updated in the background and as such users can quickly exploit the new, improved versions. SignON now supports the following spoken languages: Irish, English, Dutch and Spanish. Our machine translation component, or the intermediate language representation, initially started with only two spoken languages. We have now added other spoken languages and improved the translation quality for the pre-existing languages. The interlingual approach we have adopted allows for other languages to be easily accessed, that is, having our InterL model allows out-of-the-box translation between 20 spoken languages; however we do not provide support for them as the model has not been fine-tuned for the specific use-cases nor languages outside of the scope of SignON.[16] We are currently working on support for SLs as well. We have developed a pipeline and models for generating abstract meaning representations from text which reduces the complexity for SL synthesis. We also address the issue of low-resource data through generating synthetic gloss-to-text data and also languages built on multilingual data. On the synthesis side, the 3D animated character is built via a platform which captures the human movements and learns from them. This is a great way to improve support for existing languages as well as for integrating new languages. Furthermore, we use existing manually crafted SL data banks (or sign repositories) to bootstrap the synthesis of the avatar as well as automatically generated movements.

Many SL corpora were commonly available only for online searches. We have collected several of these corpora and made them available for download for all consortium partners. We have developed a processing tool to unify the form and organisation of these data to improve their usability. Depending on

[16] However, the SignON framework and translation-training pipeline provides the flexibility to easily extend the support to those other languages.

the licences and the informed consents of the signers, we intend to make these data sets also available to a wider public. Furthermore, we have planned two spin-off actions to collect and process available data which will be useful for the wider research community.

In addition, the major components have been developed with machine learning capabilities and the framework provides access to the corresponding interfaces so that these components can be updated with new data after deployment. Currently, an ML interface for the App is being developed to support these features.

Open Source Framework and Application Programming Interfaces (APIs):

The SignON Framework, Application and cloud platform, which features its next-generation multilingual, multi-modal user-centred functionality, will be published free and open-source, to allow the inclusion of third party functionality and modules for existing and future applications, services and languages. Through standardised APIs to the SignON services as well as to its underlying components - the ASR, SLR, InterL, NLU, and SL Synthesis - other tools and frameworks can exploit the results of this project and benefit from their functionality independent from the usage domain, objectives and technology. For many, this would translate to compliance with ISO/IEC 40500:2012.4 9.[17] This will enable the SignON Framework to become a *de-facto* tool in industry and academia.

Impact Progress Indicators from the Grant Agreement:

1. Increasing community, traffic and contributions of the open-source code released on GitHub.
2. Use of APIs in other projects or products within the lifecycle of the project.
3. Number of citations/mentions of the SignON framework or APIs.

Progress Indicators to Date:

So far, all of the code we have been developing is hosted in repositories in Github. Due to updates in licensing and pricing, we had to migrate our repositories from GitLab, where we initially hosted our Framework, Application and development repositories to Github in September 2022. We plan to release the code-base to the community after each stable version is ready. This is scheduled for 2023. The SignON platform is modular and operates through various connectivity interfaces (APIs) which allows for the easy upgrade, or replacement of such components. Through dockerising[18] components and using

[17] https://www.iso.org/standard/58625.html

[18] A docker container is a standard unit of software that packages up code and all its dependencies so an application can run quickly and reliably from one computing environment to another

external providers (e.g. acapela-group[19] for text to speech synthesis) the SignON platform is flexible, extendable and easy to test. Once released, it should be straightforward to adopt in research and industry.

It is important to note here that, while early adoption is very important for such a framework, it is also very important that the community is well informed, that expectations are managed well, and that the released code is clear, stable and easy to adopt. Balancing between these two, which depends on (i) the development and (ii) the community is a challenging task which we address jointly - developers (MAC, FINCONS) and SignON community representatives (EUD, VGTC). As such, we have planned our releases for the 3rd year of this project.

3.1.2 Additional Impacts

Greater autonomy:

Deaf signers can engage with hearing speakers immediately, without the delay (sometimes days/weeks) and administrative burden often associated with securing authorisations for providing a face to face interpreter. This also allows for more immediate engagement with the wider hearing community in low-risk scenarios. That is, SignON does not seek to replace face to face interpreting for scenarios that entail risk, e.g. legal/medical encounters. This is something that has been emphasised through our engagement with DHH communities and we are mindful of. We ought to note that SignON does not seek to replace face-to-face human interpreting when such is available, but aims to extend the repertoire of communication means.

Impact Progress Indicators:

1. Indicated in co-creation surveys at the beginning and end of the project.

Progress Indicators to Date:

In the co-creation iterations to the first prototype SignON app V1.0 reported in D1.9, the users gave the SignON mobile app a System Usability Scale (SUS) rating of 62.6 overall, which is just below the SUS threshold of acceptability of 68. Both SL users (SUS=63) and other users (SUS=62.5) rated it very similarly, but both under 68. This is to be expected as only the speech and text translations are functional, and

(https://www.docker.com/resources/what-container). Dockerisation or dockerising refers to the process of converting software into docker containers.

[19] https://www.acapela-group.com/

there are no SL recognition functions operationalised yet. However, these SUS scores indicate that the SignON app V1.0 prototype is on the right track with regard to what users want and need. The cognitive walkthrough structured formative evaluation indicates that a usable prototype app has been developed and is a good foundation for the future evolution of the SignON service for all users. They agree that it has the right look and feel. The text and speech translations are well on the way, but the SL translation functions need to be developed and be simple, and be available as soon as possible.

With new models and versions of the app being iteratively deployed, we have scheduled the next round of co-creation events and defined the SignON app demonstration and feedback required at those co-creation events, which will inform us about the progress with regards to achieving greater autonomy.

Potential to leverage SignON to supplement the work of face to face interpreters:

In cases of incidental conversation, e.g. when human interpreters require a break, but a deaf participant attending an event may wish to have a chat with another person/network or where the deaf person wishes to engage directly without working via a human interpreter, for informal conversation with a peer over a break in a meeting, for children in playground engagement (i.e.), etc..[20] SignON is going to facilitate such communication.

Impact Progress Indicators from the Grant Agreement:

1. Endorsement of the SignON application by SL interpreter services, e.g. Irish Remote Interpreting Service (IRIS)[21].

Progress Indicators to Date:

Such endorsements and adoption of SignON depend on the quality of the application and MT. As currently we have working prototypes, but no publicly accessible MT to show, it cannot go beyond what was discussed in the focus groups and round tables, which were hypothetical discussions on possible uses departing from the assumption that there is a (perfectly) working SL MT. As such we have **no observable progress** on this indicator but the attitude and perceptions of the participants in these co-creation events. For now we can only base ourselves on those hypothetical discussions during the co-creation events which indicate that deaf users suggested using the app in low-impact situations in hearing environments.

[20] Revised from the grant agreement to make it more clear within the context of this deliverable.
[21] https://slis.ie/iris/

Multiplier effect of SignON technology as a resource:

Potential to leverage SignON by users, language learners (including parents of deaf children, teachers, friends), "hearing" environments (employers, colleagues, shopkeepers, pharmacy, etc.), interpreters, deaf people, as a point of reference. For example, hearing SL learners could use the SL recognition and translation to their spoken language to assess their progress. Deaf interpreters can also use the SignON application as a tool to assist in interpreting from a spoken language to a sign language (currently they frequently use a hearing "feed" interpreter as an intermediary, or live captioning provision, which entails extra costs for the customers). As noted in the previous indicator, whether this will be the case or not, depends on the quality of the machine translation and the application, still in development and prototype releases, but also on the involved users (incl. interpreters). It boils down to whether or not deaf people decide to use the app in a specific context. The SignON application therefore interprets the spoken input into an intermediate form of sign language that subsequently serves as input for the deaf interpreters to interpret it into a more finished (more human) sign language.

Impact Progress Indicators from the Grant Agreement:

(Indicated in co-creation surveys at the beginning and end of the project (users who are learners of SL))

1. Adoption of the SignON service in education, i.e. as a possible learning tool for sign language students.

Progress Indicators to Date:

There have been none observed to date.

Machine Translation:

The project will provide access to other SL corpora and the opportunity to produce valuable research into the linguistics of these under-resourced languages with a view to developing state-of-the-art machine translation tools that have the potential to facilitate better communication between DHH and hearing people.

Impact Progress Indicators from the Grant Agreement:

1. New corpora published to open-access repositories (e.g. http://opus.nlpl.eu/).
2. Improved MT performance measured by standard evaluation metrics such as BLEU and TER as well as judged by human evaluators (deaf and hearing).

3. Validation of the machine learning technology of SignON by the academic (citation number of SignON publications) and the user community

Progress Indicators to Date:

The CorpusVGT (https://www.corpusvgt.be/)[22] was previously only available as a web-based corpus, and not as a downloadable corpus. We have obtained the data and will make it available as a single-file downloadable corpus in the CLARIN infrastructure soon. This limitation is required by the permissions given by the signers that appear in the corpus, and which stipulates it can be used for non-commercial research only. It is currently already available to the SignON consortium members.

The Corpus NGT and the ECHO[23] corpus are available from the Language Archive at Max Planck Institute, but on a file by file basis. We have collected these files in order to make them available to the consortium. Furthermore, these data have been processed and aligned by the tool described in De Sisto et al. 2022[24]. A similar issue has occurred for the BSL corpus. After investigating the BSL corpus, it has turned out that its licensing requires a substantial fee for using the BSL corpus for research. The requested GBP 80.000 made it impossible to use this corpus and we have resorted to other available resources. Based on the licences of all accessed corpora, we will make raw and processed data available, when the licence allows.

Furthermore we are in the process of processing each of these datasets to make them suitable for use in a machine learning context. Either the processed data or the scripts to process the data, or both will be made publicly available.

We are continuously experimenting with text-to-text, SL-to-text and text-to-SL MT. We have trained effective bilingual and multilingual models based on mBART[25]. We have deployed these models to the SignON framework and they are accessible via the App. After the release of the SLR model, our

[22] Van Herreweghe, M., Vermeerbergen, M., Demey, E.,De Durpel, H., and Verstraete, S. (2015). Het Corpus VGT. Een digitaal open access corpus van videos and annotaties van Vlaamse Gebarentaal, ontwikkeld aan de Universiteit Gent i.s.m. KU Leuven.

[23] Brugman, H., Crasborn, O., and Russel, A. (2004).Collaborative annotation of sign language data with peer-to-peer technology.InProceedings of theFourth International Conference on Language Resources and Evaluation (LREC'04),

[24] De Sisto, M., Vandeghinste, V., Egea Gómez, S., De Coster, M., Shterionov, D. & Saggion, H., (2022). Challenges with Sign Language Datasets for Sign Language Recognition and Translation, Proceedings of the 13th International Conference on Language Resources and Evaluation (LREC). Marseille, France

[25] Liu, Y., Gu, J., Goyal, N., Li, X., Edunov, S., Ghazvininejad, M., ... & Zettlemoyer, L. (2020). Multilingual denoising pre-training for neural machine translation. Transactions of the Association for Computational Linguistics, 8, 726-742.

immediate next task is to assess the machine translation performance for Sign-to-text. Based on our findings we will work to improve these models. It is worth noting that now, by connecting SLR and InterL, the complexity has increased and the optimisation problem involves multiple factors.

Clear social justice and equality agenda that we help to achieve via SignON:

The SignON project will promote fairness and equality through the inclusive approach of developing the SignON service that brings the DHH and hearing communities in an open dialogue mediated via the experts in our consortium. Through better understanding of communities' individual needs as well as the requirements to facilitate fair communication between them, SignON will deliver a communication service that will reduce social inequality and inclusion imbalance.

Impact Progress Indicators from the Grant Agreement:

1. Transparent reiterative cycles of discussion practising active engagement with and between DHH and hearing communities across the life of SignON.

Progress Indicators to Date:

The users' feedback reported in the user requirements reports (D1.3 and D1.7) and the technical requirements reports (D1.4 and D1.9) clearly indicate that open dialogue and inclusive involvement of the DHH users and deaf communities in developing the SignON app, framework and services will reduce social inequality and inclusion imbalance. We have practised open communication with the potential users and designed and developed the SignON app and framework according to their requirements (as defined in deliverables D1.4 and D1.9). We will continue this practice so as to maximise the impact with respect to social justice and equality. Furthermore, through our SignON REC we ensure that all members of the consortium follow the established ethics and privacy guidelines and principles defined in WP9. Currently we have conducted the following co-creation events: Four focus groups in four countries (IRL, NL, BE, ES), one survey (BE) one workshop (Ireland) and one roundtable (NL), which have allowed members of the potential user communities to openly communicate about the impact of technology. Furthermore, we plan to expand the reach of these events and increase our inclusiveness through events with special attention to diversity and inclusion of the different minorities within the deaf communities.

Education:

The learning from the SignON project will offer a new resource to facilitate the instruction of life skills among vulnerable deaf people with limited literacy skills which DCU, as a third-level educational institute, would look to initiate. It will also allow further development of certain SL grammar books and will lead to a significant speed up in the annotation of SL corpora. In this way, strong points of two independent groups can be joined to come up with something completely unprecedented, allowing easy access to syntactically annotated SL corpora which is something that KU Leuven wishes to exploit.

Impact Progress Indicators from the Grant Agreement:

1. Schools, universities and organisations representing and supporting deaf people present the platform as a way to engage with deaf people in a more human way reducing barriers to communication and promoting inclusion
2. Endorsement of the platform by the deaf communities, interpreters and hard-of-hearing people in the pilot sites

Progress Indicators to Date:

While at present there is no observed impact of the SignON services or the work conducted within the project on education, the fact that SignON raises awareness about sign languages has led to members of the consortium organising sign language classes. In particular, the team from TiU has organised NGT classes for the researchers of the university (currently commencing level 3). The teams at DCU and TCD have been exploring ways to expand on the existing co-creation approaches within SignON. To supplement existing focus groups and roundtables, TCD and DCU collaborated with the Education and Public Engagement team at ADAPT to pilot their Think-In method within SignON. This involved multiple structured engagements with the community to raise awareness about SignON, to provide information on the technology, and to facilitate discussions on how this technology might impact the Deaf community, and society more broadly. To explore the potential of contemporary science communication practice for co-creation within SignON, the team at DCU adapted an art-science method originally developed by TRACES as part of SISCODE. With support from Science Foundation Ireland, DCU collaborated with Deaf theatre makers to produce a public Shakespeare performance that integrated elements of machine vision and machine translation. This was followed by an audience discussion to find out if this immersive experience might spark new perspectives and feedback from the community. We are in the process of evaluating and analysing these approaches, and we hope that it will inform future approaches to engagement and informal education.

Lowering gender bias:

SignON will bring a service that lowers gender bias: (i) ASR, SLR and SL Synthesis will be adaptable to the user's preferences, including gender; (ii) the recognition, translation and synthesis tools will be developed following SOTA research on reducing gender bias, e.g. using gender-balanced data; (iii) participants in co-creation events will be selected in a gender-balanced way; (iv) with the 60/40 distribution of male to female institution leaders in our consortium the SignON project will be a positive example of unbiased synergy.

Impact Progress Indicators from the Grant Agreement:

1. Ethics guidelines are observed;
2. The co-creation approach and the measurements for reducing gender bias we develop within the SignON project are adopted by other European initiatives and projects.

Progress Indicators to Date:

The SignON research ethics committee (REC) was established at the beginning of the project. The SignON REC has compiled an ethical guidelines document (D9.1) and ensures these are observed across the life of the project. The committee supports the process towards securing ethical approvals for research relating to the project, and assists in responding to any ethical queries arising at project level. The SignON REC has supported five applications for ethics approval across 2022 (one to be submitted in January 2023).

OpenSource tools for web-based interactive 3D graphics:

UPF-GTI will be able to extend the capabilities of the OpenSource tools for web-based interactive 3D graphics it has been offering to the community through GitHub, improving the capability to edit, animate and render virtual characters with much higher quality.

Impact Progress Indicators from the Grant Agreement:

1. Increased traffic to the GitHub repository;
2. Endorsement of the 3D virtual signers and puppeteering system by other European projects, researchers and industry stakeholders.

Progress Indicators to Date:

The two web applications hosting i) the automatic generator of animations (for the moment only for the manual features (MFs)), and ii) the 3D scene with the Behaviour Realiser moving the avatar, can be found

in the two different GitHub repositories. The first application is a new contribution in the estimation of animations from a single-camera frame, which also includes an editor of animations and an exporter into BVH animation files. The second application holds the 3D scene that will be displayed to the final users, and is real-time linked to the Database of MoCap animations. It is planned to integrate in the main mobile application in the next stages of the project (since it was delayed due to other setbacks in the planned input data). The mentioned repositories can be found in the following urls:

i) https://github.com/upf-gti/SignON-editor

ii) https://github.com/upf-gti/SignON-realizer .

3.1.3 Impact to consortium partners

The collaboration between the partners of the SignON consortium will expand their professional network and will create new avenues for knowledge transfer. It will have a common as well as an individual positive impact. Here is a rundown of the impact that our consortium partners have seen to date:

DCU: Our role in SignON will strengthen the position of DCU as a leader in MT. This project is contributing greatly to the many projects within DCU working towards educational equality for DHH students, including the ISL STEM Glossary project and the Bachelor of Education (ISL). This project will have practical use to the Deaf staff and students studying and working in the Faculty of Education in DCU. Our role in the co-creation work package has given us the opportunity to explore novel approaches to community engagement and engaged research practice, including ADAPT's Think-In approach, as well as an art-science approach initially developed by TRACES as part of SISCODE.

FINC: The SignON project allows FINC to continue its effort and commitment in the development of accessibility services for current and prospective customers. Thanks to the leading role in WP2 and to the interaction with partners, FINC had the chance to be involved in the co-creation of the SignON App, framework and services, which helped FINC be more engaged in the requirements and needs of DHH people and deaf communities. Moreover, as a result of developing the SignON Framework with the SignON Orchestrator, APIs and services (T2.2 and T2.4), FINC had the chance to deepen its understanding of the specific technical requirements needed to support software modules in the processing pipeline and MT training, enhancing FINC capabilities related to the development of applications with a strong focus on accessibility services for all citizens.

INT: As a consequence of participating in the SignON project, INT has enlarged its portfolio of linguistic services with respect to Inclusion. This fact was well appreciated by the external audit committee in 2021. INT has also released several sign language datasets for download (VGT corpus, BeCoS corpus[26]), presented these on conferences and published about them in scientific journals, and announced these in their newsletter and on social media, which impacts how the public perceives the institute, not merely a service centre for theoretical linguists and lexicographers (which it used to be), but as an institute with societal relevance. Working with sign language corpora has brought the institute expertise in working with video data. Setting up the infrastructure for SignON has brought expertise about setting up powerful GPU servers. INT has also developed a model for punctuation insertion for Dutch and for multilingual data, initiated as a requirement for the BeCoS corpus. This model is applicable to ASR output and is made publicly available on HuggingFace. It has had several thousand downloads.

UPV/EHU: UPV/EHU's previous experience was restricted to the written form of spoken languages. SignON has allowed them to extend this experience to multimodal frameworks that include a wider range of languages (both spoken and signed languages), thus opening up new possibilities for research and collaboration. In fact, thanks to the experience gained in the project, they have obtained a new national project (PID2021-123988OB-C31) in which, in collaboration with other Spanish institutions, they will work on the translation of languages with very few resources, including Spanish Sign Language.

MAC: As a result of developing and publishing the prototype SignON mobile app V1.0 (D2.6), producing the first and second “Technical Requirements & User Research (UX design) Reports” (D1.4 & D1.9) and documenting the required SignON app feedback at co-creation events, MAC staff are now more engaged and better aware of the requirements and wishes of DHH people and deaf communities. The MAC team are also now more experienced in the use of structured formative users’ evaluation and a user driven co-creation iterative technology design approach based on User Experience (UX) design thinking techniques. Working on WP2, MAC staff are now more aware of the technical options and potential of the open SignON app, framework and services. In task T2.4, the MAC team are learning the requirements of MT training, from developing the SignON ML app to train the SignON MT. Researching D6.6 “SignON Market Analysis“ and D6.7 “First SignON Sustainable Exploitation, Innovation and IPR

[26] Vincent Vandeghinste, Bob Van Dyck, Mathieu De Coster, Maud Goddefroy, Joni Dambre (2022). Belgian Covid-19 Sign language corpus. A corpus for training Sign Language Recognition and Translation. *CLIN Journal.*

Plans“, the MAC Team have become much more aware of the SL/MT market situation, competitors and exploitation potential of SignON.

UPF: The contributions of UPF to this project can be divided according to the different research groups that are participating in SingON: TALN and GTI.

On the side of UPF-TALN, we have extended the experiments reported in the first project year and disseminated the findings in different forums. Particularly, we experimented with different deep learning blocks to include linguistic information in the Text-to-Gloss translation task. The results and findings produced in these experiments were published in the NLDB2022 conference. Additionally, we researched how to improve model performances under very-low-resourced languages, the case of LSE. We proposed a text augmentation method and a pretraining strategy to boost LSE gloss translation. These contributions were published in the LoResMT workshop at COLING 2022.

Regarding data acquisition and processing, UPF-TALN contribution is two-folded: (1) we designed a ELAN corpus processing framework to extract useful data from SL corpus and generating data ready to be fed into deep learning models; and (2) we have scrawled a LSE dictionary including videos and text information from https://fundacioncnse-dilse.org. The processing framework produced in (1) was published in LREC 2022 and we are still improving the processing pipeline for efficiency.
The UPF-TALN latest experiments include proof of concepts in finetuning mBART to include SLs at text level and experiments in NGT Sign Classification using OpenPose and MediaPipe features.

On the side of UPF-GTI, we have improved the quality of our previous virtual characters in order to meet the requirements of this project. There have been improvements specifically in the mesh and the textures, so the avatar has more quality overall, but mainly in some key parts (such as hands and face) in order to give more realism to the final render and support a wider range of facial expressions. Aside from that, we have developed two applications (explained in “OpenSource tools for web-based interactive 3D graphics” from Section 3.1.2): i) one of them involving novel approaches in the problem of “synthetic animations”, and contributing to the problem of “scarcity of suitable datasets”; ii) and the other one providing a light-weight interactive 3D scene (with a realizer module to reproduce BML orders procedurally) intended to be used in the SignON mobile application.

TUDublin: As a result of participation within the SignON project TU Dublin has gained extensive insight into the varying SOTA approaches to SLMT, providing us with a toolkit to extend our own potential in solving this very important and challenging problem leveraging not only a linguistically informed approach but also opening up an opportunity to tackle this problem using a cutting edge hybrid approach. The project has very much broadened our links and collaboration with colleagues working within this interdisciplinary field, together with our understanding of the importance of ethics and engagement with DHH communities allowing us to leverage this understanding and inform our research and development. We have presented our work in progress at both national and international level including, IRAAL 2021 and the international AMTA conference 2022 among others

TCD: Following on from participation in SignON, the Centre for Deaf Studies has expanded its' range of multidisciplinary engagement, learning about the key tools, processes and procedures that underpin the work of MTSL. Engagement in SignON has opened up our engagement with hard of hearing and deafened people, cohorts that CDS does usually engage with given our primary focus on sign languages. We are also learning about and documenting how attitudes to MTSL are changing, informed by the pandemic and how this impacted on online engagement, and this informs how we are interrogating how DHH communities view the ethics of MTSL and other technologies that impact on DHH peoples' lives. Additionally, we are learning more about how ISL is changing, e.g. through our analysis of newly developed glossaries for political terms, STEM concepts, Gender Based Violence terms, and terms linked to the Black Lives Matter movement. We have presented on work in progress at a number of national and international conferences and seminars including IRAAL 2021, LREC 2022, ISGS 2022, etc. as well as referencing SignON at a Think Tank on the future of sign language interpreter education (Humboldt University, Berlin 1-2 December 2022). Through leading WP9, we have learned a great deal about the varied approaches to ethics in our consortium's institutions and, from engaging with colleagues from INT, about CLARIN and data curation and data management plans.

VRT: As a result of participating in the SignON project, and especially in the feedback that came from the co-creation workshops with the deaf and hard of hearing communities, the public media broadcaster VRT receives more knowledge into what impactful changes it can do to create more inclusive content. These learnings are being shared with producers and content creators in the organisation via internal

presentations and workshops. At the technological side: the target communities have asked for more subtitled content; this has led to VRT investing in 2 new funded AI based subtitling projects. Meanwhile, a second item was mentioned regarding the avatars. People request life like avatars when used in professional broadcasts. VRT innovation is researching the future possibilities to bring SignON avatars to Unreal's MetaHuman.

UGent: Ghent University, in particular the IDLab-AIRO research group, researches SLR in cooperation with members of the Flemish deaf community and SL linguistics researchers. Thanks to the SignON project, we have been able to shift our focus towards multilingual SL recognition and SL translation. This provides us with new insights into SLs and how machine learning algorithms can learn their common elements, while at the same time improving the performance of language-specific SLR models. The interaction with industrial partners has allowed IDLab-AIRO to broaden our knowledge on the deployment of SL recognition applications, which will enable new collaborations with the deaf community in Flanders. From the interaction with SL linguists, we continue to learn more about SLs, which will be beneficial to our SL processing research. The collection of new large datasets in SignON will allow us to evaluate our technology in large-scale SLT experiments.

VGTC: Thanks to the co-creation events, VGTC has learned that there is not only the negative resistance of the deaf signers to sign language technology. There are also positive expectations about the SignON project. The evolution of sign language technology also had an impact on the operation of VGTC. Since it was not one of the important priorities in language planning for Flemish Sign Language in policy of VGTC, this domain was ignored before. Now it becomes an important part of our operation and planning for the coming years. Also the confidence in (hearing) researchers has grown thanks to the co-creation events that have had a real impact on our co-design and co-development. This gives VGTC a better understanding of sign language technology. Within the communication of VGTC we received various questions about this matter, which we can better answer and/or argue now, also based on scientific publications.

UCD: The team at UCD has gained hugely from working with low-resource sign languages, prompting a great deal of innovation in developing deep learning methods specifically for low resource scenarios which can differ greatly from standard, data hungry deep learning techniques. The insights gained from the research have enabled us to publish at an international venue (SLTAT 2022) and have motivated work

on other resource constrained machine learning tasks, specifically synthetic data generation for ISL and domain-specific gesture recognition within the wider group at UCD. The co-creation aspect of SignOn has also given the team at UCD the opportunity to understand the wide array of considerations that must be taken into account when developing SLR technology and has motivated the group to consider the end users and stakeholders at every level of development. In fact, some members of the team have even gone on to take modules in Irish Sign Language in order to deepen their linguistic understanding of sign languages. Furthermore the diverse expertise of SignOn has enabled the members of UCD to reach out to linguists in the consortium with ease, allowing us to create more nuanced technical advancements in the SLR component of the SignOn application.

RU: At a local level, RU strengthened the involvement in ASR technology of the SL group in the Faculty of Arts of RU through the SignON project. From an international perspective, it helped RU obtaining international funding from CLARIN ERIC to set up a special CLARIN Resource Family page for Sign Language resources (see https://www.clarin.eu/resource-families/sign-language-resources). Through this funding we realised intensive collaboration with other CLARIN Knowledge Centers focusing on SL in Europe. Further, RU was at the initiative of the hospitality use case recordings for the next phase of the SignON project. This brought us more involvement and collaboration with the SL oriented partners in the project, and with communities and networks of DHH communities. Finally, the work on tuning ASR systems specific for DHH speakers puts RU at the forefront of inclusive ASR development.

TaalUnie: For Taalunie as the language policy organisation in the Dutch language area the SignON project has been particularly important to draw attention to the two sign languages in the area, VGT and NGT, and the role human language technologies can play to increase accessibility, inclusivity and equality. Although the NGT partners were not initially involved in the project, we have found ways of unlocking the NGT data available at the TV and radio companies through the Sound and Vision Institute to the benefit of the SignON project and we have managed to set up a fruitful cooperation with the NGT centre. This will increase the acceptability and future impact of the SignON applications.

KU Leuven: As far as KU Leuven's Taalgroep Vlaamse Gebarentaal (Flemish Sign Language group) / MIDI Research Group is concerned, the SignON project has led to a further extension of our multidisciplinary collaboration (in particular a collaboration with computational linguists & AI experts) and an introduction to the key tools, processes and procedures that underpin the work of MT and SLR. This collaboration

opens up possibilities for (current and future) linguistic research, e.g. through a facilitation of the annotation of corpus data. The already existing relationship with the VGTC has been strengthened, while our participation in SignON has opened up engagement with the Flemish hard of hearing and deafened people, cohorts that we usually do not collaborate with, given our primary focus on Flemish Sign Language. Further, within the framework of SignON, we learned how to organise a large-scale written survey targeting hearing, deaf and hard-of-hearing audiences. Our participation in WP1 and particularly in WP9 taught us a lot about the ethical aspects of MT and AI applications/research related to sign languages. As part of our SignON research activities, we are now initiating a study to investigate whether and in which way certain grammatical mechanisms, tendencies and language use in Flemish Sign Language are influenced when deaf signers communicate with each other through live and postponed video-mediated communication (2D situation) instead of face to face communication (3D situation).

TiU: Tilburg University has been involved in two main streams of tasks. First it is the management task, together with DCU and second, the coordination and the (local) research activities on WP4. Through the first set of tasks, Dimitar Shterionov and Mirella De Sisto have strengthened their management and coordination skills, become more involved and aware of REC procedures and requirements, have expanded their collaboration network and learned a lot about the deaf community. Through the NGT classes organised by Mirella De Sisto, Tilburg University staff members are given the opportunity to learn NGT. On a university level, the SignON project has increased the visibility of the university, the school and the department and further supported the research and development for reducing bias and increasing equality. TiU has organised two out of the three co-creation events in the Netherlands building a stronger collaboration with the NGTC. Through the SignON collaboration network we have identified additional data needs and prepared additional research activities to address them. This gives us the opportunity to work in smaller, focused teams on an important topic. In 2022, TiU organised the 32nd edition of the CLIN conference which had a SL track which included 6 presentations on SL-related topics such as SLT, SLR, segmentation and avatar synthesis.

EUD: The EUD works for equality in the public and private life of deaf people. SLs play a key role in achieving this equality and in the accessibility of deaf persons and human rights in general. To this end, the SignON project is an excellent opportunity to get truly involved in SLs, on both a human and technical side. EUD took the opportunity to include users in a co-creation process and to represent users as best as possible while collaborating with technical partners on the tremendous possibilities that the

technology can offer. The inclusion of end-users was achieved by organising co-creation events in the scope of WP1. This way, valuable feedback could be gathered at the very start of the SignON project and throughout its life. The inclusion of user feedback from the co-creation events and its analysis was necessary to address the needs of deaf people in line with the Convention on the Rights of Persons with Disabilities. The inclusion of EUD in an EU project of this scale, makes sure that the needs and priorities of the end-users of the SignON project are put to the forefront from the beginning. The importance of this must be underlined, as in the past very few projects included deaf users in the development from start to finish. This did not fit in with the “Nothing about us without us'' principle. EUD sincerely believes that the SignON project sets an example and has achieved transparency in its work process. The involvement in a project of this kind also deepens the internal knowledge at EUD regarding issues related to sign language technologies, which can be used for lobbying and/or future projects.

3.2 Measures to Maximise Impact

At the beginning of the SignON project a communication and dissemination plan (D6.1) was prepared. This plan ensures proper visibility, accessibility and promotion of SignON and its results during the project lifecycle and after its completion. We target 5 groups of users: (i) deaf communities (signers), (ii) DHH people (non-signers), (iii) hearing people, (iv) SL interpreters, and (v) scientific community. To reach all these groups in a timely and appropriate manner, the consortium relies on the extensive network already established by SignON participants representing the community and industry partners as well as the significant international reach of the External Advisory Board. We have already engaged and informed these groups through (i) scientific and popular publications; (ii) conference, workshop and seminar presentations; (iii) wide-ranging media coverage (newsletters, social media posts, press releases). In 2022 we have presented different aspects of the project, including the avatar, the translation capabilities, etc., in our co-creation events, as well as at the workshop at the EU Parliament. The application and framework have been tested and validated by internal users. While the initial idea was to demonstrate the SignON framework and app in front of open audiences, we have decided to limit the outreach to selected members initially because we want to increase the awareness before releasing a product that is undergoing continuous research and development. Having learned from previous experiences, presenting a product in development could be misunderstood or misinterpreted and lead to failure in future adoption.

The SignON website, www.signon-project.eu, is the dissemination hub for our project. For the wider outreach, the content on our website is available in English and in International Sign (IS); the updated overview about SignON 2022 is translated to all national sign languages (soon to be published) Furthermore, the website and its content is developed by deaf professionals who understand the communication needs of deaf communities.

In 2022, SignON participated in the organisation of three conferences - SLTAT2022, WMT 2022 - Shared task on SL translation and CLIN32. SLTAT and WMT were organised by EASIER[27] and SignON partners.

All consortium partners are actively involved in the communication and dissemination activities, which are being led by VGTC (communication and dissemination) and MAC (exploitation). The ADAPT Centre/DCU is the managing body of this project and provides legal and privacy advice and support.

The consortium tracks market trends to identify business strategies and roadmaps for the exploitation of Intellectual Property (IP) arising from the project's work. Deliverable D1.1 *Case studies and evidence analysis* presents our first analysis of use-cases, market and prior research; we continue to do so on a regular basis. In doing so, we will be able to discover new market trends and opportunities to incorporate new approaches while supporting efforts to demonstrate SignON's appeal and potential profitability to potential investors.

4. Impact Progress to Date

The aim of the SignON service is to bridge a communication gap between deaf communities, people who are hard of hearing (including people with cochlear implants) as well as hearing people (including family members and friends of deaf people), i.e. to contribute to the communication repertoire between these user groups. It therefore aims to connect industry partners and policy makers, and to promote inclusion of deaf communities by providing an open source and scalable solution for real-time translation between SL, text and speech. With the expected wide uptake of the technology, SignON will make a significant contribution to its users and our target communities.

SignON also impacts academic research, industry innovation and business progression through its SOTA advancements in SL linguistics, SLR, NLU, ASR (including atypical speech), MT and 3D avatar technology, and their fundamentals. When openly released, our platform and its constituent tools will offer an

[27] https://www.project-easier.eu/

alternative vision for the governance and values of SL translation and will be publicly available to end-users in the form of a mobile app for real-time translation. In the first year of the project, SignON was mainly on an exploratory mission - to identify the needs and gaps, potential technologies and existing research in order to establish a new vision for translation between different EU sign and spoken languages through a service co-created by the users. In the second year of this project, we developed most of the necessary components through robust research activities, built models for recognition, NLU, translation, synthesis and integrated those components and models in an effective framework accessible via a lightweight application. We continue to analyse and document the (hitherto under-documented) languages we are working with. The SLR component has been recently integrated in the framework and the SLS component will be integrated in the framework in early 2023.

As noted in Section 3.2., the target audiences will be engaged and informed via (i) co-creation events, (ii) scientific and popular publications, (iii) organisation of and participation in conference, workshop and seminars, (iv) wide-ranging media coverage and (v) prototype presentations and demonstrations.

- Co-creation events: WP1 organised several interviews, either with multiple participants (a focus group), or with just one person. Interviews were held with participants from Flanders, the Netherlands, Ireland and Spain. Furthermore, physical co-creation events were organised; a workshop took place in Dublin, Ireland and a roundtable took place in Amersfoort, the Netherlands. To ensure a diverse group of participants, people from different ages and genders were invited. The impact is such that the potential (DHH) users felt listened to during these co-creation events, which resulted in a strong willingness to follow-up further stages of SignON development (e.g. to test each component of the SignON services if they have the opportunity). The impact is the positive view on the SignON project, thanks to being aware that deaf people are involved from the start this time and that we will actually listen to them. Currently other co-creation events are being arranged to take place in the third year of the SignON project.
- The SignON Consortium was invited by Ádám Kósa, a deaf Member of the European Parliament, to present the SignON project at the European Parliament in Brussels. We took the opportunity to also invite representatives of the National Associations of the Deaf from the project countries, to attend the presentation of the project in the European Parliament and afterwards hold a separate workshop to strengthen the bonds with the national deaf communities even further.
- Scientific and popular publications: The list of work that spans from the SignON project and published in academic and non-academic venues is:

- 2022:
  - Mirella De Sisto, Vincent Vandeghinste, Santiago Egea Gómez, Mathieu De Coster, Dimitar Shterionov and Horacio Saggion (2022). Challenges with Sign Language Datasets for Sign Language Recognition and Translation. Proceedings of the 13th International Conference on Language Resources and Evaluation (LREC). Marseille, France.
  - Vincent Vandeghinste, Bob Van Dyck, Mathieu De Coster, Maud Goddefroy and Joni Dambre (2022). BeCoS corpus: Belgian Covid-19 Sign language corpus. A corpus for training Sign Language Recognition and Translation. Computational Linguistics in the Netherlands Journal. Vol. 12
  - Vincent Vandeghinste and Tim Van de Cruys (2022). Voorgetrainde Grote Taalmodellen en Automatische Vertaling. Dixit. Tijdschrift voor Taaltechnologie.
  - Dimitar Shterionov, Mirella De Sisto, Vincent Vandeghinste, Aoife Brady, Mathieu De Coster, Lorraine Leeson, Josep Blat, Frankie Picron, Marcello Paolo Scipioni, Aditya Parikh, Louis ten Bosch, John O'Flaherty, Joni Dambre, and Jorn Rijckaert (2022). Sign Language Translation: Ongoing Development, Challenges and Innovations in the SignON Project. In Proceedings of the 23rd Annual Conference of the European Association for Machine Translation, pages 325–326, Ghent, Belgium. European Association for Machine Translation.
  - Matthieu De Coster and Joni Dambre (2022). Leveraging Frozen Pretrained Written Language Models for Neural Sign Language Translation. In Information volume 13. p. 220.
  - Aditya Parikh, Louis ten Bosch, Henk van den Heuvel, Cristian Tejedor-García (2022). Design principles of an Automatic Speech Recognition functionality in a user-centric signed and spoken language translation system. In Computational Linguistics in the Netherlands Conference (CLIN32), TIlburg, the Netherlands.
  - Ruth Holmes, Ellen Rushe, Frank Fowley, Anthony Ventresque (2022). Improving Signer Independent Sign Language Recognition for Low Resource Languages. In Proceedings of the 7th International Workshop on Sign Language Translation and Avatar Technology: The Junction of the Visual and the Textual: Challenges and Perspectives (SLTAT 2022), pp. 45-52. Marseille, France.

- Rachel Moiselle and Lorraine Leeson (2022). Language Planning in Action: Depiction as a Driver of New Terminology in Irish Sign Language. In Proceedings of the 10th Workshop on the Representation and Processing of Sign Languages (sign-lang@LREC 2022), pages 139–143 Language Resources and Evaluation Conference (LREC 2022), Marseille, 20-25 June 2022.
- Mathias Müller, Sarah Ebling, Michèle Berger, Richard Bowden, Annelies Braffort, Necati Cihan Camgöz, Cristina España-Bonet Roman Grundkiewicz, Eleftherios Avramidis, Alessia Battisti, Zifan Jiang, Oscar Koller, Amit Moryossef, Regula Perrollaz, Sabine Reinhard, Annette Rios, Dimitar Shterionov, Sandra Sidler-Miserez, Katja Tissi, Davy Van Landuyt (2022). Findings of the First WMT Shared Task on Sign Language Translation (WMT-SLT22). In Proceedings of the Seventh Conference on Machine Translation (WMT), pages 744–772, Abu Dhabi, December 7–8, 2022.
- Vincent Vandeghinste, Mirella De Sisto, Dimitar Shterionov, AoifeBrady, Mathieu De Coster, Loraine Leeson, Josep Blat, Frankie Picron, Marcello Scipioni, Aditya Parikh, Louis Ten Bosch, John O'Flaherty, Joni Dambre, and Jorn Rijckaert (2022). Progress, Challenges and Innovations of the SignON Project. Proceedings of the Computational Linguistics in the Netherlands Conference (CLIN32).
- Egea Gómez, S., Chiruzzo, L., McGill, E., Saggion, H. (2022). Linguistically Enhanced Text to Sign Gloss Machine Translation. In: Rosso, P., Basile, V., Martínez, R., Métais, E., Meziane, F. (eds) Natural Language Processing and Information Systems. NLDB 2022. Lecture Notes in Computer Science, vol 13286. Springer, Cham.
- Coppin, J., De Coster, M., & Dambre, J. (2022). Attention analysis of a sign language recognition task on the AUTSL dataset. Computational Linguistics in the Netherlands (CLIN 32), Abstracts. Presented at the 32nd Meeting of Computational Linguistics in The Netherlands (CLIN 32), Tilburg, the Netherlands.
- Ineke Schuurman, Vincent Vandeghinste, Caro Brosens, Margot Janssens and Thierry Declerck (2022) WordNets? SignNets! (abstract) Presented at the 32nd Meeting of Computational Linguistics in The Netherlands (CLIN 32), Tilburg, the Netherlands.

- Ineke Schuurman, Thierry Declerck, Caro Brosens, Margot Janssens, Vincent Vandeghinste and Bram Vanroy (accepted) Are there just Wordnets or also SignNets? Global WordNet Conference 2023, Donostia-San Sebastian. Basque Country (Spain)
- Luis Chiruzzo, Euan McGill, Santiago Egea Gómez, Horacio Saggion. Translating Spanish into Spanish Sign Language: Combining Rules and Data-driven Approaches. LoResMT@COLING 2022: 75-83 (https://aclanthology.org/2022.loresmt-1.10.pdf)
- Shaun O'Boyle, Elizabeth Mathews, Davy Van Landuyt, Frankie Picron, Rehana Omardeen, Lorraine Leeson, Rachel Moiselle, Aoife Brady, Jorn Rijckaert, Caro Brosens, Anthony Ventresque, Ellen Rushe, Irene Murtagh, Andy Way, Dimitar Shterionov (2022). Using co-creation to develop sign language machine translation technologies (Abstract). Poster presentation at SCI:COM 2022, Dublin, Ireland.
- Frank Fowley, Ellen Rushe and Anthony Ventresque (2022). "A Data Augmentation and Pre-processing Technique for Sign Language Fingerspelling Recognition". Irish Machine Vision and Image Processing Conference (IMVIP)
- Frank Fowley and Anthony Ventresque (2021). "Sign Language Fingerspelling Recognition using Synthetic Data". 29th Irish Conference on Artificial Intelligence and Cognitive Science (AICS)

- 2021:
  - Mathieu De Coster, Mieke Van Herreweghe and Joni Dambre, Isolated Sign Recognition from RGB Video using Pose Flow and Self-Attention, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2021
  - Dimitar Shterionov, Vincent Vandeghinste, Horacio Saggion, Josep Blat, Mathieu De Coster, Joni Dambre, Henk van den Heuvel, Irene Murtagh, Lorraine Leeson, and Ineke Schuurman, The SignON project: a Sign Language Translation Framework, 31st Meeting of Computational Linguistics in the Netherlands (CLIN31). 2021

- H. Saggion, D. Shterionov, G. Labaka, T. Van de Cruys, V. Vandeghinste and J. Blat, SignON: Bridging the gap between Sign and Spoken Languages, XXXVII Spanish Society for Natural Language Processing conference (SEPLN2021). 2021
- Dimitar Shterionov, John O'Flaherty, Marcello Paolo Scipioni, Matteo Villa, Edward Keane and, Marco Giovanelli, Early-stage development of the SignON app and open framework - challenges and opportunities, Proceedings of the 18th Machine Translation Summit (MTSummit2021). 2021
- Mirella De Sisto, Dimitar Shterionov, Irene Murtagh, Myriam Vermeerbergen and Lorraine Leeson, Defining meaningful units. Challenges in sign segmentation and segment-meaning mapping, Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL). 2021.
- Mathieu De Coster, Karel D'Oosterlinck, Marija Pizurica, Paloma Rabaey, Severine Verlinden, Mieke Van Herreweghe and Joni Dambre, Frozen Pretrained Transformers for Neural Sign Language Translation, Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL). 2021.
- Santiago Egea Gómez, Euan McGill and Horacio Saggion. Syntax-aware Transformers for Neural Machine Translation: The Case of Text to Sign Gloss Translation, Proceedings of the 14th Workshop on Building And Using Comparable Corpora. (RANLP 2021), 2021.

- Organisation of and participation in conferences, workshops, symposia and other forums. In 2022, SignON was also involved in organising the Sign Language Translation and Avatar Technology Workshop (SLTAT 2022), the shared task on machine translation for sign and spoken languages within the WMT 2022 conference and CLIN32, including the special track on sign languages. Furthermore, SignON had presentations at CLIN32
- Wide-ranging media coverage: SignON content is disseminated via its channels in Facebook (SignONEU), LinkedIn (SignON (EU project)) and Twitter (@SignONEU). A great success is the SignON encyclopaedia which presents videos containing information about sign language, avatars, deafness, machine translation and other interesting and related to the project topics. The SignON website as a dissemination hub.

Since the major update of the website and its publication, and the launch of the social media channels of SignON, our target groups are aware of our project. At the start of the project the general view was that the potential user groups still experience SignON as something not necessarily relevant to them personally. However, through the different communication and dissemination activities and co-creation events, these views are changing. It is obvious that through mutual understanding the project is evolving into a technological hub for all potential users. Not only "they know that the SignON project exists and what its purpose is" they now have an idea of the added value and the drawbacks of such a project and can make an informed decision when and whether to use the technology; they are more open to contributing and expressing their opinion about use-cases and the future of sign language translation and the use of 3D virtual signers (i.e. avatars) . We have now created and populated an FAQ-page on the SignON website with questions and answers that originated from the various communication and co-creation activities. Furthermore, SignON has built collaboration bridges with other projects, researchers and organisations, including Dr. Maartje De Meulder, who presented along with the SignON consortium at the EU Parliament, a partnership with NGTC on organising co-creation events in the Netherlands, strengthening the relations with the EASIER project and others.

The dissemination activities that have taken place to date as part of the SignON project are further elaborated on in deliverable D6.3: Second Annual Reports on Communication and Dissemination Activities which will be submitted to the EU Portal at the end of December 2022.

The plans for exploitation routes within and outside the SignON consortium have been outlined and discussed in deliverable D6.7: "First SignON Sustainable Exploitation, Innovation and IPR plans", which has been submitted at the end of the first year of the project. The final SignON exploitation and innovation plan will address our plans beyond the lifecycle of the project and will be submitted in 2023.

5. Conclusions

This report presents the impact achieved in the second year of the SignON project. It covers scientific, technical and societal impacts. It aligns the current impact with what was defined in the grant agreement and shows the progression from the first year of the project. The report focuses on the communication and dissemination activities conducted in order to maximise the impact.

In 2023 we plan to deliver 29 deliverables. . These will lead to the following impacts:

- To have the second version of the SignON App released with SLR and SLS functionality which will impact the direct stakeholder communities. Reciprocally their feedback will provide new insights and impact the continuous development of the SignON application and service;
- To have a separate SignON ML App for recording SL and spoken language to train the SIgnON MT
- The release of new research and development in the fields of SLR, ASR, MT, SLT, SL linguistics and avatars which will impact both academic and industry landscapes;
- Through continuous and open communication with our stakeholder and target communities, we will expand our information exchange, mutual understanding and acknowledgement of differences, while reducing biases and (potentially) inequality.