Sign Language Translation Mobile Application and Open Communications Framework

Deliverable 1.1: Case Studies and Evidence Analysis

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-FT-57-2020
Project Start Date: January 1st 2021

Deliverable Information
Title: D1.1 Case Studies and Evidence Analysis
Work Package: 1 - Co-creation and User Response
Lead Beneficiary: TCD
Due Date: 31/03/2021
Revision Number: V0.3
Author: L. Leeson (TCD)
Dissemination Level: Public
Deliverable Type: Report

Revision History

Version #	Implemented by	Revision Date	Description of changes
V0.1	TCD	15/03/2021	First Draft
V0.2	TCD	26/03/2021	Partners' contributions
V0.3	TCD	30/03/2021	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	D1.1	Aoife Brady	DCU	30/03/2021
V0.1	D1.1	Vincent Vandeghinste	INT	30/03/2021
V0.2	D1.1	John O'Flaherty	MAC	30/03/2021
V0.1	D1.1	Horacio Saggion	UPF	30/03/2021
V0.2	D1.1	Gregg Young	VRT	30/03/2021
V0.2	D1.1	Dimitar Shterionov	TiU	30/03/2021
V0.2	D1.1	Frankie Picron	EUD	30/03/2021

Acronyms

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

Acronym	Definition
MT	Machine Translation
ASL	American Sign Language

FAIR	Findability, Accessibility, Interoperability, and Reuse
DMP	Data Management Plan
ISL	Irish Sign Language
IS	International Sign

Table of Contents

1. Overview

This report draws on available literature that describes, documents and interrogates best practices with respect to sign language machine translation (MT) with a particular emphasis on the adoption of application-based technologies and services.

Firstly, we address deaf and hard of hearing stakeholders' views towards MT, drawing on what is limited literature. This sets the scene for work in progress to document current attitudes and perceptions of European deaf signers vis-a-vis a range of MT approaches including sign languages, a theme we will report on later in the SignON project.

Here, we also consider data available to our project and outline key considerations relating to data protection during and into the post-project deployment phase of the SignON project.

A number of projects that entail sign language translation have been supported over the past decade or so. Here, we summarise a sample to indicate current state of the art. Against this backdrop, we will explore deaf community responses to these developments, identify other key challenges (e.g. range of data available, limited documentation of sign languages, insufficient sign language repertoires in many teams, and crucially, lack of engagement with Deaf communities and limited involvement of deaf scientists on many (most) teams.

2. Background

A primary objective of the SignON project is to create a service that translates between sign and verbal languages, facilitating new resource generation over time, which in turn will further improve the service. This, however, is not a simple endeavour.

Sign languages differ from territory to territory, for example: in Ireland there are 5,000 deaf Irish Sign Language (ISL) signers; in the UK around 87,000 deaf signers use British Sign Language (BSL); in Flanders, Belgium some 5,000 deaf people use Flemish Sign Language (VGT); and in France approximately 100,000 are deaf native French Sign Language (LSF) signers; in 2019 there were approximately 60,000 users of sign language of the Netherlands (NGT); it is estimated that there are over 100,000 speakers of Spanish sign language (LSE).

Sign languages are not simply visual representations of verbal languages. They are independent natural languages expressed in the visual-gestural modality. Their grammars are expressed using the hands, torso and face of the signer, often in complex constructions that depart quite significantly from the patterns of production we usually think about with respect to verbal languages. We might think about spoken languages as beads on a string; sign languages can be thought of as layer cake like, with the potential for clausal meaning presented via several channels simultaneously, with the combined gestalt representing the intended meaning. Sign languages have what have been described as both an established lexicon and a productive lexicon. The established lexicon is the set of fairly fixed vocabulary that has a relatively one-to-one mapping to a word (e.g. signs exist for concepts like *mother*, *father*, *home, etc.*). The productive lexicon comprises a set of language-specific handshapes that can combine with a wide range of movements, orientations of the hand, and locations in the signing space to create dynamic meaning. These are accompanied by facial gestures (mouth gestures, eye-gaze patterning, brow-raises/brow-furrows, ...) and may represent clauses or sentences encoding a particular character perspective on an event or encoding particular discourse focus, and are particularly challenging for verbal language to sign language machine translation and vice versa, not least because of the ubiquitous 'depiction' strategies.

Before we turn to outline the key challenges outlined in the literature that are associated with machine translation of sign languages and the adoption of technologies that have been developed to date, we briefly outline the current state of the art in this field.

3. What we can Learn from Projects on Machine Sign Language Translation to Date

The goal of SignON is to develop a free, open-source service and framework for conversion between video (capturing and understanding our target sign languages), audio (for speech, including atypical speech) and text, translating between sign and verbal languages, delivered to its users via an easy to use mobile application.

The project is driven by a focused set of use-cases tailored towards deaf and hard of hearing communities. To date, the use-cases of MT solutions for sign languages are few and very narrow; often the work on MT for sign languages aims at general solutions, i.e. solutions that could apply in any use-case and for any sign language, which is typical for many works on text-to-text MT.[1] Such use-cases involve, e.g. translation of digital TV content; translation of bus information, or translation of children's books into a range of sign languages. Indeed, Jantunen et al. (2021) report that we can categorise the technologies that are relevant to the automated translation of sign languages into two overall groupings: (1) those that require wearable technologies (gloves, accelerometers, etc.) (e.g. see Cheng et al. 2013, Nan et al. (2014), Gupta et al. 2016, Kiliboz, N. Ç., Güdükbay, U. 2015) and (2) those that seek to capture signers' hand gestures from a distance via camera or sensor-based tracking (e.g. see Molchanov et al. 2015, Lee and Park 2009, Bao et al. 2017, Liang et al., Mantecón et al. 2014, 2016, 2019; Cui et al. 2019). What we should also point out here is that there is a distinction between translating from a sign language and translation to a sign language, with each direction requiring different technological pipelines. However, this is not something that we drill down into in this report. It will, however, be dealt with in other SignON project deliverables.

At this point we ought to note some approaches that motivate and support the research and development activities scheduled in SignON related to the translation process. They can broadly be categorised as end-to-end versus pipelined. A software pipeline typically refers to a sequence of components that transform an input into a desired output by independently applying the processing steps within each component on the input or an intermediate representation of it. An end-to-end approach typically applies a joint set of transformations on the input to generate the output, without any essential intermediate steps or representations.

- Pipelines: Typically there are two types of MT considered in the context of a sign language:
    - Translating a sign language into a verbal language which is typically implemented by the following (high-level) pipeline a) recognizing sign language input and generating of an intermediate representation based on, e.g. glosses[2] or HamNoSys[3]; and b) translating from this intermediate representation into a textual representation using an MT system, which in turns can be either a pipelined approach, e.g. phrase-based statistical MT with moses[4] or an end-to-end, e.g. a neural MT model developed with a tool such as OpenNMT[5]

[1] We ought to note that there is a large subfield in MT that aims at domain-specific solutions. However, it is typical for MT to start from a general-domain and then narrow it down to a domain-specific solution.
[2] https://www.lifeprint.com/asl101/topics/gloss.htm
[3] Hanke, Thomas. "HamNoSys-representing sign language data in language resources and language processing contexts." *LREC*. Vol. 4. 2004.
[4] www.statmt.org/moses/
[5] https://opennmt.net/

  - Translating a verbal language into a sign language is a different process as it does require the synthesis of a visual representation of a signer. This is done through the use of automated or semi-automated 3D actors (often referred to as avatars). The typical automated avatar pipeline involves a) the translation or encoding of verbal language into a format that defines spatio-temporal dependencies of different anchor points of a 3D actor such as SigML[6] and b) the rendering of such an actor, e.g. JASigning[7]. In the context of semi-automated pipeline, typically text is a) translated to glosses or other intermediate representation which is associated with a pre-recorded avatar movements, then b) a message in a sign language is generated as the concatenation (and often 3D transition smoothing) of the different pre-recorded avatar movements. An example of such pre-recorded avatar movements is the Simax avatar.[8]
- End-to-end: advances in current data-driven machine learning and deep learning have motivated the development of models that can jointly recognize and translate (sign-to-text), such as the recent works of Yin, Kayo, and Jesse Read[9] and of Camgoz et al[10]; or translate and synthesize (test-to-sign). Regarding the latter case, an innovative work that was recently published and is developed within the project Content4All is Text2Sign by Stoll et al.[11] where instead of 3D avatars, sign language is generated by the use of neural models based on generative-adversarial networks that aim to mimic real (human) signers.

The end-to-end approaches bypass a lot of issues related to error propagation between the different components in a pipelined architecture. However, they tend to require a lot of data as a model needs to train on a large number of examples, e.g. signed messages and text in the target language for an end-to-end sign-to-text translation. The SignON project aims to address numerous translation directions, involving 3 different modalities. The data requirements are very high which, together with the fact that sign language data is scarce, makes an end-to-end approach within SignON impractical. That is why we opted for a pipeline approach that involves recognition, translation (through an intermediate representation) and synthesis of a 3D avatar.

[6] https://vh.cmp.uea.ac.uk/index.php/SiGML
[7] https://vh.cmp.uea.ac.uk/index.php/JASigning
[8] https://simax.media/?lang=en
[9] Yin, K., & Read, J. (2020, December). Better sign language translation with stmc-transformer. In Proceedings of the 28th International Conference on Computational Linguistics (pp. 5975-5989).
[10] Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10023-10033).
[11] Stoll, S., Camgoz, N. C., Hadfield, S., & Bowden, R. (2020). Text2Sign: towards sign language production using neural machine translation and generative adversarial networks. International Journal of Computer Vision, 128(4), 891-908.

Furthermore, SignON seeks to incorporate sophisticated machine learning capabilities that allow (i) learning new sign and verbal languages; (ii) style-, domain- and user-adaptation and (iii) automatic error correction, based on user feedback and overcoming obstacles related to resource availability.

In Section 5, we consider state of the art with respect to the key challenges facing those seeking to develop machine translation driven responses to access. First, however, we consider what we know of deaf community views on technological responses to translation requirements.

4. Deaf Community Perspectives

There has been very little empirical work published around the experiences/preferences of deaf signers or hard of hearing people vis-à-vis machine translation software and/or devices, something that SignON wishes to address. What has been published around the response of deaf communities to avatars and other technologies tends to be published not in academic fora but in mainstream media, if it is reported on at all. This signals a disconnect between those engaged in work to develop and support machine translation with sign language users. However, what has been documented reflects quite a negative response. For example, Erard (2017) notes that wearable technologies, such as signing gloves, which claim to translate American Sign Language (or indeed, any sign language) overlook the intricacies of the language and the needs of signers.

Some part of the problem clearly relates to how technologies are presented and reported – there is a clear need to ensure that those who are presenting on research and reporting in the media are mindful of their ethical obligations around the work they do and the implications of their claims. Another part of the problem relates to how out of kilter the research goals of those working on technological responses are to the lived experience of deaf communities. The risk of excluding deaf community members or deaf users of these products from the process of development, is that inventors are likely to continue creating devices that offend the very group they say they want to help. As Thad Starner, director of the Contextual Computing Group at the Georgia Institute of Technology notes, “To do this work, the first rule you have to teach yourself is that you are not your user.” (Erard, 2017).

Given the lack of inclusion of deaf signing communities in processes designed to develop tech-driven responses, it is not surprising deaf community responses to such technological inventions have been quite negative. Several examples illustrate why this is the case.

In 2017, O'Connor et al. (2017) published a paper on a wearable glove designed by engineers at the University of California, San Diego that set out to translate the ASL alphabet to written English on a computer or smartphone. This received significant media attention which hailed it as a glove that 'translates sign language' (Medgadget editors, 2017) and that, in turn, prompted a significant negative response from the American Deaf Community, not least because the capacity to recognise and translate some of the fingerspelled letters of a signed alphabet is quite different from translating a language. Indeed, it is akin to being able to recognise a small number of sounds from a spoken language but not being able to recognise those sounds when they occur as part of a sequence of sounds.

Another wearable glove project, SignAloud, which was designed by students at the University of Washington, and won the Lemelson-MIT Student Prize was similarly problematic. SignAloud secured coverage in the media (e.g. see NPR[12] and Bustle[13]) but was also critiqued heavily by linguists. For example, Dr. Agnus Grieve Smith's blogpost, "Ten reasons why sign-to-speech is not going to happen any time soon"[14] opens with:

> "It's that time again! A bunch of *really eager* computer scientists have a prototype that will translate sign language to speech! They've got a *really cool video* that you *just gotta see*! They win an award! (from a panel that includes no signers or linguists). Technology news sites go wild! (without interviewing any linguists, and sometimes without even interviewing any deaf people). …and we computational sign linguists, who have been through this over and over, every year or two, just *facepalm*."

Grieve Smith is not alone in her response.

Erard (2017) cites the ASL programme director at the University of Washington, Lance Forshay who says that he was surprised and "…felt somehow betrayed because they [the SignAloud students] obviously didn't check with the Deaf community or even check with the ASL program teachers to make sure that they are representing our language appropriately".

---

12 https://www.npr.org/sections/alltechconsidered/2016/05/17/478244421/these-gloves-offer-a-modern-twist-on-sign-language?sc=17&f=3?sc=17&f=3&t=1616878698973
13 https://www.bustle.com/articles/157062-these-gloves-translate-sign-language-into-text-speech-in-real-time
14 http://grieve-smith.com/blog/2016/04/ten-reasons-why-sign-to-speech-is-not-going-to-be-practical-any-time-soon/

Professor Carol Padden, a highly regarded American deaf sign linguist, noted that the SignAloud gloves, like all similar sign translation gloves developed to date, misconstrue the nature of ASL and other sign languages. Going beyond the challenges of the signing glove, Padden notes that:

> "One challenge for machines is the complexity of ASL and other sign languages. Signs don't appear like clearly delineated beads on a string; they bleed into one another in a process that linguists call "coarticulation" (where, for instance, a hand shape in one sign anticipates the shape or location of the following sign; this happens in words in spoken languages, too, where sounds can take on characteristics of adjacent ones). Another problem is the lack of large data sets of people signing that can be used to train machine-learning algorithms." (cited in Erard, 2017).

Table 1. SignON Use Cases

In reviewing the State of the Art (SOTA), Bragg et al. (2019) point to the following areas and consider the major challenges that exist in each domain.

- Datasets
- Recognition and Computer Vision
- Modeling and Natural Language Processing
- Avatars and Computer Graphics
- UI/UX Design

In the following sections, we outline the issues identified in turn.

5.4.1 Datasets

Existing sign language datasets vary in terms of data content, the discourse function of such content, the number of signers, and critically, signer identity and status (e.g. first/second language user; deaf signer, adult child of deaf adults (CODA), non-deaf ("hearing") signer, interpreter, etc.).

Bragg et al. (2019) note that the source of data is important, both in terms of content and signer identity. For example, some corpora consist of fluent signers paid to create target corpus content; some consist of public videos such as those found on YouTube or other sites; others again may include interpreted content (e.g. news channels output that is interpreted into a sign language).

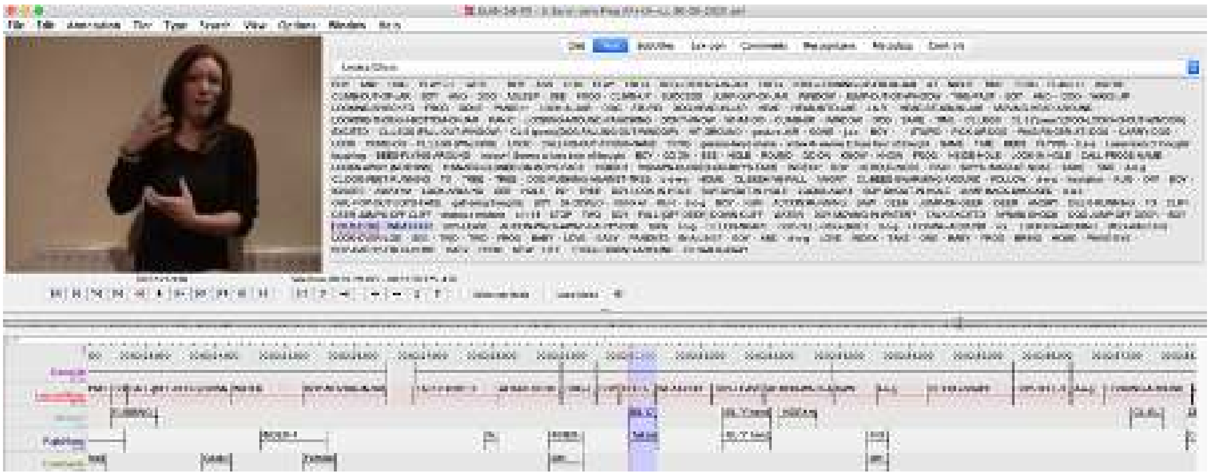
Bragg et al. note that the bulk of curated data available is in American Sign Language, but they point to other data sets for a range of sign languages that vary in size from about 100-2000 distinct signs (Matthes et al., 2012). Further, the format of such data varies significantly. Some data sets (e.g. those created using motion-capture) are curated and drawn on for the purpose of developing signing avatars. Bragg et al. (2019) present an overview of American public corpora used in sign language recognition projects on a regular basis, which we present here to illustrate the variety and extent of source types (Table 2).

Dataset	Vocabulary	Signers	Signer-independent	Videos	Continuous	Real-life
Purdue RVL-SLLL ASL [65]	104	14	no	2,576	yes	no
RWTH Boston 104 [124]	104	3	no	201	yes	no
Video-Based CSL [54]	178	50	no	25,000	yes	no
Signum [118]	465	(24 train, 1 test) - 25	yes	15,075	yes	no
MS-ASL [62]	1,000	(165 train, 37 dev, 20 test) - 222	yes	25,513	no	yes
RWTH Phoenix [43]	1,081	9	no	6,841	yes	yes
RWTH Phoenix SI5 [74]	1,081	(8 train, 1 test) - 9	yes	4,667	yes	yes
Devisign [22]	2,000	8	no	24,000	no	no

Table 2: Popular public corpora of (American) sign language video commonly used for sign language recognition (Bragg et al. 2019, p. 20).

Consideration of data source, content function, and signer identity are issues that we have raised to date in our SignON project team meetings (re: WP3), identifying the need to work, in the first instance, with publicly available corpora featuring deaf signers (e.g. the VGT corpus, the ISL corpus, the NGT corpus) and other data such as sign content that is publicly available but not yet digitally annotated. The nature of annotation is problematised too: how data is annotated varies significantly. However, in recent years, most sign language corpora that have been developed are annotated using the Max Planck Institute's ELAN software[15], which allows for recourse to the source signed content, as in the example below, courtesy of the Signs of Ireland corpus (Frog Story, Sarah Jane). Another consideration is the data format which needs to become consistent between different corpora in order for a unified processing approach, as the one intended by SignON, can be possible.

[15] https://archive.mpi.nl/tla/elan



However, the SignON team will have to develop an intermediate representation to encode both sign and verbal language to consequently support the machine translation process. This intermediate representation, or InterL, needs to cope with the complexities of sign languages - complexities that are just in process of being described linguistically (and informed, we add, by work in gesture studies). (e.g. Stokoe & Wilcox 1995, Wilcox 2007).

Another challenge for our team is the extent of appropriate available data. The SignOn team has noted that the bulk of data available via broadcast media sources is, in fact, translated - sometimes by deaf translators or interpreters, and sometimes by hearing interpreters. This risks increasing levels of influence from the spoken/written source language onto the target sign language output, which is not ideal. Thus, in addition to considering available data, we must also consider the function of texts and the discourse types available. Finally, we must consider the signer's status; our goal is to leverage deaf signer created content insofar as possible. We note that while there is a good deal of publicly available data featuring sign language interpreters, most hearing interpreters are second language users of a sign language which impacts on proficiency vis-a-vis native signers. These are issues that we address in WP3.

5.4.2 Recognition and Computer Vision

Bragg et al. note that, from as far back as 1983, glove-based approaches to sign language recognition have been used to circumvent computer vision problems involved in recognizing signs from video (p. 19), but as we note elsewhere in this report, they are highly problematic, and widely disliked by deaf communities for many reasons. They note that non-intrusive vision-based sign language recognition is the current dominant approach, and have the benefit of minimising inconvenience to the signer but introduce complex computer vision problems. They point to the seminal work of Tamura and

Kawasaki, who built a system to recognize 10 isolated signs of Japanese Sign Language (Tamura & Kawasaki, 1988). However, they note that many other systems have also focused on identifying individual signs (Grobel & Assan, 1997; Lang et al., 2012). On this point, they outline the challenges of continuous recognition of signs but note that this is a more realistic problem than the identification of single signs, given that it is "…confounded by epenthesis effects (insertion of extra features into signs), coarticulation (the ending of one sign affecting the start of the next), and spontaneous sign production (which may include slang, non-uniform speed, etc.)." (p.19). They note that addressing the 3-D nature of sign languages in such contexts thus requires a range of approaches including use of depth cameras, use of multiple cameras, or triangulation for 3D reconstruction. They also note that while recent advances in machine learning have improved SOTA computer vision approaches, lack of sufficient training data currently limits the use of modern Artificial Intelligence (AI) techniques in this problem space. Bragg et al. also note that while automatic recognition systems are in the process of transitioning from small artificial vocabularies to larger real-life vocabularies, recognition systems currently achieve circa 42.8% letter accuracy (e.g. see Shi et al.'s (2018) work on a real-life fingerspelling dataset). In addition to fingerspelling recognition, two other sub-tasks of sign language recognition can be discerned: isolated and continuous sign (language) recognition. As mentioned by Bragg et al., isolated sign recognition research often utilises datasets collected in a lab setting specifically for machine learning. Few works focus on isolated sign recognition with co-articulation. Albanie et al. (2020) collect a dataset of 1064 signs from British Sign Language with co-articulation. Their best model reportedly can recognize signs with 70.38% accuracy. De Coster et al. (2020) perform isolated sign recognition on a dataset built from the corpus of Flemish Sign Language (VGT), also featuring co-articulation. They report obtaining 74.70% accuracy on 100 classes. Continuous sign recognition typically is benchmarked on the RWTH-PHOENIX-Weather dataset (Forster et al., 2012), collected from weather broadcasts. This dataset contains 1080 different signs from German Sign Language. The current state of the art on this dataset is a word error rate of 20.7 (Zhou et al., 2020). While research is moving towards these real-life vocabularies, there is still considerable room for improvement in the proposed methods for sign recognition.

5.4.3 Modeling and Natural Language Processing

We have noted earlier that sign languages are typically under documented. They are minority languages and, in comparison to their spoken language counterparts, very little machine translation or natural language processing (NLP) work has focused on them. Bragg et al. (2019, p.20) report that:

> "While recognition handles the problem of identifying words or signs from complex signals (audio or video), MT and NLP typically address problems of processing language that has already been identified. These methods expect annotated data as input, which for spoken languages is commonly text (e.g., books, newspapers, or scraped text from the internet). Translation between spoken and signed languages (and vice versa) also typically requires intermediary representations of the languages that are computationally compatible."

They report on methods that are applied in the field as including both predefined representations (e.g. Veale et al. 1998; Elliot et al. 2000; Zhao et al., 2000, Davydov & Lozynska 2017), which are compatible with grammatical translation rules and those that use deep learning or neural networks which learn model features to address the challenges presented (e.g. Camgoz et al. 2018; Camgoz et al. 2020; Koller et al. 2016, 2018). Such methods, they add, have been used for recognition combined with translation.

5.4.4 Avatars and Computer Graphics

In Section 3, we presented a broad overview of the state of the art for the field of technologically mediated translation of sign languages to/from spoken languages. Here, we add that sign language avatars have been used to provide accessible content in a sign language to deaf and hard of hearing communities who prefer to access information in a sign language. They also have the potential to open up access to information for those who are not fluent in the written language of the jurisdiction they are in (e.g. Bragg et al. 2019, Jantunen et al. 2021, Mathews & O'Donnell 2018). Further, we know that deaf communities report that they prefer human signers - and ideally want to communicate with their interlocutor via a sign language (Kyle, Sutherland and Stockley 2012, Napier & Kidd, 2013; Napier, Sabolcec et al., 2014). It has also been shown that deaf people in many countries report that they prefer deaf interpreters, particularly for content broadcast on television or online (Kyle 2007). It is the case that there simply are not enough signers or interpreters in the world to facilitate access to the range of contexts that contemporary deaf communities wish to engage within, and there are economic constraints that operate too. With this in mind, we can appreciate that Bragg et al. (2019) suggest that avatars may be more appropriate when automatic generation of content is desirable, as in the case of websites with unstable content. Bragg et al. (2019, 20-21) add that, at present, "The state-of-the-art in avatar generation is not fully automated; all parts of current pipelines currently require human intervention to generate smooth, coherent signing avatars."

SignON addresses this challenge by drawing on partners who are experts with experience in avatar creation, informed by cross-disciplinary engagement and deaf community stakeholder feedback loops to help to exponentially improve avatar acceptability over the life of the project. We aim to address this in two ways: (i) develop an avatar with sophisticated features that have been presented to and agreed by the deaf community; (ii) integrate machine learning capabilities in our system that will allow the system to be improved by human professionals. With respect to (ii) users would have the opportunity to teach the system how to respond better to their needs.

We are also extremely aware of the concerns of deaf communities (to date, stated anecdotally) that there is a risk that state bodies may seek to defund much needed and hard won provisions for face-to-face human interpreting by assuming that technological solutions meet their legal obligations to provide reasonable accommodations, without understanding the limitations of such technologies or the consequences of assuming they are a sufficient replacement for human interpreter provision. SignON emphasises that **we do not seek to replace human interpreting provision in any way**, but rather seek to open up access, securing greater access and participation than is currently possible.

5.4.5 UI/UX Design

Bragg et al. (2019) report that "The state-of-the-art of sign language output in user interfaces primarily centers around systems that use sign language video or animation content (e.g., computer-generated human avatars) to display information content." (p.21). Projects using UI/UX design include those which seek to provide sign language animation content to supplement text content for sign language users, and projects that seek to present links to ASL dictionary resources to provide on-demand terminology definitions (e.g. Hariharan et al. 2018). They go on to note that research focused on the design of interactive systems using sign language recognition technologies has focused on the creation of useful applications despite the limited accuracy and coverage of current technology for this task. Such applications have included tools for students learning ASL, either young children (e.g., Zafrulla et al. 2011) or older students who are provided with feedback as to whether their signing is accurate (Huenerfauth et al. 2017). Further, they point out that while several projects and industry efforts have sought to create tools that can recognize full phrases of ASL to provide communication assistance, few systems are robust enough for real-world deployment or use.

6. Bragg et al.'s Call to Action and the SignON Response

Against the backdrop outlined in section 5, we turn to consider Bragg et al.'s (2019) call to action, which entails five key points which can also function as a framework of reference against which we can and will evaluate our SignON process and our products.

Call 1: Involve Deaf team members throughout projects. They note that "Deaf involvement and leadership are crucial for designing systems that are useful to users, respecting Deaf ownership of sign languages, and securing adoption." (Bragg et al. 2019, p. 23).

SignON is predicated on a co-creation model which places deaf users at the heart of our process.

Call 2: Focus on real-world applications. They note that sign language processing is appropriate for specific domains, and that the technology has limitations. They argue that datasets, algorithms, interfaces, and overall systems should be built to serve real-world use cases, and account for real-world constraints. (Bragg et al. 2019, p.23).

Following from our co-creation model, SignON use cases will be determined by and fine tuned recursively by our deaf community partners and wider network of deaf stakeholders.

Call 3: Develop user-interface guidelines for sign language systems. Bragg et al. note that as sign language processing is still developing as a field, there is a lack of systematic understanding of how people interact with it. They argue that guidelines and error metrics for effective system design would support the creation of consistently effective interfaces. (Bragg et al. 2019, p. 24).

Call 4: Create larger, more representative, public video datasets. The authors argue that large datasets with diverse signers are essential for training software to perform well for diverse users. They note that public availability is important for spurring developments, and for ensuring that the Deaf community has equal ownership. (Bragg et al. 2019, p.24).

This is, broadly speaking, outside the scope of the SignON project, but is a call that we will echo in our work. There is clearly a need for greater documentation of European sign languages as a precursor to facilitating better machine translation outcomes. Such datasets also play other roles for deaf communities - they support linguistically under-resourced communities in the digital sphere - less than 1% of content on the internet is in a sign language (Leeson et al. 2015).

Call 5: Standardize the annotation system and develop software for annotation support. Working with sign language data sets, Bragg et al. note that annotations are essential to training recognition systems, providing inputs to NLP and MT software, and generating signing avatars. Standardization would support data sharing, expand software compatibility, and help control quality. Annotation support would, therefore, help improve accuracy, reliability, and cost. (Bragg et al. 2019, p.25)

While this call is outside our SignON remit, we are cognizant of how little annotated sign language content is available, and the range of approaches that have been taken in annotating the sign language data for the languages that we work with. We will, through our processes, document the challenges that this brings, and support the call for the development of software that helps to automatise elements of annotation processes for sign languages.

7. Research Integrity and Data Management Concerns

In this section, we turn to consider some ethical considerations that relate to working with deaf communities and then turn to consider data protection issues as they pertain to sign language content.

7. 1 Ethical considerations

Wallwork (2002, p.21) argues that "the partnership ideal usefully suggests that our research ethics itself needs to be jointly negotiated and constructed among mutually respectful participants, willing to be changed through dialogue about how to cooperate in joint undertakings." This is an ideal we take seriously in the SignON consortium. Yet, deaf communities are, as we have noted, minority language communities. They have experienced suppression of their languages, and, frequently, exclusion on the basis of deafness, despite the existence of legal instruments that consider language rights and/or disability status (e.g. Ladd 2003, Wheatley & Pabsch 2012, Tupi 2019). Much of this has arisen at the hands of hearing, non-signing policy makers, educators, often informed by hearing, non-signing researchers. In such circumstances, the ideals of partnership, of co-creation of outcomes, of "nothing about us without us" have not been embraced.

In section 5, we noted that there has traditionally been limited engagement of those working on machine translation and other technologically driven projects with deaf communities. Additionally, we have seen that there have been few examples of deaf professionals engaged in projects such as SignON. This has led to a situation of mistrust, and feelings of being excluded from the funding streams that support such projects, and with that, excluded from the process of research that will, above all else, impact on deaf communities. Anecdotally, there is a sense in deaf communities that funders are trusting hearing researchers who know very little about deaf communities over members of deaf communities themselves, and supporting projects that are not typically considered to be key priorities by members of deaf communities.

In any research context, trust between researchers and participants is essential - clearly, given the context we are working in on the SignON project, and given our partnership make-up, we are dedicated to ensuring that we build trust through transparent, engaged, inclusive processes. We adopt the ALLEA European Code of Conduct for Research Integrity (2017) principles of reliability, honesty, respect and accountability in all we do. Additionally, we draw on the literature that has been published that offers ethical guidance to researchers working with deaf communities, and leverage this in our project deliverable D9.1 - Ethical Guidelines and Protocols document (Due in Month 6).

Here, we note that Harris et al. (2009, p.109) advise that when working with deaf communities, we must consider the importance of respect, beneficence, and justice. They say:

> "…respect is defined in terms of the cultural norms of interaction within the Sign Language community and throughout the hearing and D/deaf worlds. Beneficence is defined in terms of the promotion of human rights and increased social justice. An explicit connection is made between the process and outcomes of research and furtherance of a social justice agenda."

They go on to propose that a Terms of Reference for Academic Research and Publications for researchers working with signing communities is necessary, and outline a Sign Language Communities Terms of Reference (SLCTR) (Table 3) that builds on the Indigenous Terms of Reference (ITR) (Osborne & McPhee, 2000).

1. The authority for the construction of meanings and knowledge within the Sign Language community rests with the community's members.
2. Investigators should acknowledge that Sign Language community members have the right to have those things that they value to be fully considered in all interactions.

3. Investigators should take into account the worldviews of the Sign Language community in all negotiations or dealings that impact on the community's members.
4. In the application of Sign Language communities' terms of reference, investigators should recognize the diverse experiences, understandings, and way of life (in sign language societies) that reflect their contemporary cultures.
5. Investigators should ensure that the views and perceptions of the critical reference group (the sign language group) is reflected in any process of validating and evaluating the extent to which Sign Language communities' terms of reference have been taken into account.
6. Investigators should negotiate within and among sign language groups to establish appropriate processes to consider and determine the criteria for deciding how to meet cultural imperatives, social needs, and priorities.

Table 3: Sign Language Communities' Terms of Reference Principles (Harris et al, 2009: 115) (Adapted from ITR)

The SignON team adopts these principles in our work, and we will embed these in our approach to ethical engagement, folding these into the terms of reference for our Ethics Committee (WP9).

7.2 Data Management

In section 5, we noted the challenges of securing adequate amounts of data in sign languages which is sufficiently prepared for the purposes of the work that those engaged in machine translation work require. While our project will initially focus on pre-existing data sets such as digital sign language corpora and publicly available content, we have also given due consideration to how we will handle and curate new data over the life of SignON and beyond.

Pre-existing data sets are shared under the terms of the SignON Consortium Agreement, which all partners have signed up to.

New data that will be collected will require a two-stage process of ethics clearance. That is, when we seek to collect data (e.g. from participants in focus groups, or in creating additional data sets to supplement existing corpora), partners will prepare an ethics application for their home institution, or, if they are a non-university partner, an application for research ethics approval will be submitted via the coordinating partner institution, Dublin City University. Before submitting their application to their institutional research ethics committee, the application will be reviewed by the SignON Ethics Committee, who will ensure that the key principles outlined in section 7.1 and in D9.1 (to follow, M9) are embedded in our research processes and practices.

Some general principles apply:

- All SignON researchers will be required to confirm that they have received training regarding their obligations around GDPR.
- Collection of personal data will be minimised insofar as possible.
- Informed consent will be secured from participants in all SignON data collection processes.
- All participants in any data collection process will have access to information in plain language and/or a signed language, as is their preference.
- Information that is video recorded will be used only for the purposes that are advised, and kept only as long as necessary (e.g. focus group data that is video recorded will be later transcribed and pseudo-anonymised). However, we note that some data collection may enrich existing corpora and should, assuming appropriate permissions from participants are in place, be curated as FAIR[16] materials.
- SignON will create a Data Management Plan (DMP) and, as required by institutional research ethics processes, data management processes will be articulated in each new application for ethical approval prepared. In line with the Science Europe Practical Guide to the International Alignment of Research Data Management[17] we will consider our DMP under the following headings: (i) Data Description and Collection; (ii) Documentation and Data Quality; (iii) Storage and Backup during the research process; (iv) Legal and Ethical Requirements, Codes of Conduct; and (v) Data Sharing and Long-Term Preservation.
- Primary responsibility for observing good practice in the use, storage, retention and preservation of data sits with the individual SignON researcher, supported by the SignON consortium.
- Any research data that the SignON consortium collects will be recorded in a clear and accurate format. Particular attention will be paid to the completeness, integrity and security of our records.
- Any research data we collect will be stored in secure and accessible form and will be retained for a length of time in accordance with institutional, funder and/or publisher requirements (also considering FAIR usage). This information will be listed in documentation shared with participants engaging in data collection processes across the life of the SignON project.
- SignON researchers will publish results and interpretations of our research in an open, honest, transparent and accurate manner, and respect confidentiality of data or findings when legitimately required to do so. We also note that given the nature of sign language

[16] https://www.go-fair.org/fair-principles/
[17] https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

materials, the image of the participant is an essential vehicle of the linguistic message. In this, we will explore whether participants who consent to their images being used in publications/presentations/subsequent teaching and learning contexts wish to be assigned pseudonyms or if they wish to be acknowledged and named, something that many deaf participants have requested in our experience (e.g. see Leeson and Saeed's (2012) foreword acknowledging and naming (with permission) the contributors to the ISL[18] corpus).

[18] In this document, and overall through the SignON project, we use the acronym 'ISL' to refer to Irish Sign Language; to refer to International Sign, we use IS. See table of abbreviations in the beginning of the document.

5. Key Challenges

Extrapolating from the literature and media reports that do exist, we can pinpoint a number of challenges, which can be assigned to the following general categories:

5.1 Lack of Engagement with Deaf Communities

A key issue that emerges is the fundamental lack of engagement with deaf communities with regard to their priorities in terms of linguistic access by many (perhaps most) projects to date. A corollary of this appears to be the lack of understanding on the part of those engaged in technology-driven projects, including machine translation projects, of the lived experience of deaf communities. The day to day reality for deaf people is that they typically have to engage in significant recursive engagement with policy makers and service providers to ensure that their languages are legally recognised, and following from that, that appropriately skilled, qualified, human interpreters are provided in a broad range of settings (e.g. educational, medical, legal, cultural, broadcast, religious, and social settings) (De Clerck & Pinxten 2012; de Wit et al., 2020; DeMeulder & Haualand, 2019; Kermit et al., 2014; Kyle, 2007; Leeson et al., 2020; Leeson et al., 2014; Leeson & Vermeerbergen, 2012; Napier & Barker, 2004; Napier & Haug, 2016; Napier & Kidd, 2013; Napier & Leeson, 2016; Napier & Spencer, 2008; Nilsson et al., 2013; Ofcom, 2007; Pabsch & Söderqvist, 2015; Smeijers & Pfau, 2009; Vermeerbergen et al. 2012; Wehrmeyer, 2015).

This extends to the effort required on the part of deaf communities to secure interpreting provision even in the most pressing emergency situations. For example, in Britain, deaf people led the call for a case to be taken against the British government for its lack of provision of interpreting for COVID-19 briefings, fuelled by a social media campaign, #WheresTheInterpreter (Rose, 2020)[#]. Even before the pandemic, other emergencies led to communities turning to social media to insist upon the provision of appropriate access to information during emergencies (Leeson, 2019; McKee, 2014). The need for greater consideration for the communication access needs of deaf communities around the world led to publication of a joint document on this matter from the World Federation of the Deaf and the World Association of Sign Language Interpreters (World Association of Sign Language Interpreters (WASLI) & World Federation of The Deaf (WFD), 2015), one that is regularly cited on social media platforms like Twitter.

5.2 Limited engagement of deaf signing communities and/or fluent signers on MT Projects

There are exceptionally few deaf people engaged on teams working on technological projects focused on sign languages. Indeed, there are few fluent signers, hearing or deaf, working on such projects. The absence of anything approaching a critical mass of deaf people on such projects means that the cultural and linguistic capital that deaf communities can bring to technical projects is absent. Further, the linguistic competence required to plan, execute and deliver on such projects is often insufficient to facilitate stated project goals. This impacts on the quality of outputs produced. Too often, the teams presenting avatar mediated machine translated content don't know how linguistically limited their end product is because of this essential gap in linguistic competence. There is also a significant risk that by not engaging with deaf communities, and not appreciating how deaf communities view approaches adopted, that great offence will be caused, which, in turn, may further influence perceptions in deaf communities about the goals/intentions/potentiality of work on machine translation and associated technical domains in a negative way. But it goes beyond this – there is also the significant risk that hearing teams - and their funders - will invest in creating 'solutions' that are quite at odds with the priorities of deaf communities. In such instances, we find an intersection between what we might consider to be deaf and hearing ontologies (i.e. different ways of conceptualising and categorising 'deafness' versus 'hearingness' – e.g. see Ladd 2003; Young, Napier, & Oram 2020), and with issues that can be considered as research integrity matters (a point we return to later).

A case in point is the book chapter, "Implementation of Hand Gesture Recognition System to Aid Deaf-Dumb People" (Ghule & Chavaan, 2021) published (and subsequently promised for retraction[#]) by Springer in 2021 (though we note it is still currently available online). Led by Dr. Julie A. Hochgesang from Gallaudet University's Department of Linguistics, over 100 linguists, Deaf studies scholars and community members co-signed a letter to the publisher. In the letter, the co-signatories note that they were "appalled and dismayed" by the chapter, noting that the work was marred by a lack of understanding of the field, pejorative and outdated language and other problems.

They note that the publication raises significant ethical issues and point to clear guidelines that have been developed by those in the sign linguistics community in recent years to underpin such work. Key principles invoked are that "…work of this nature must, at a minimum, be done in collaboration with the communities who stand to be most impacted by this work. Better yet, this work should be led by members of these communities" (Hochgesang, 2021). The letter adds that authors and editors of any paper, in any field, have an obligation to ensure that relevant literature has been consulted, including (we add) the literature on Deaf studies and sign linguistics, which was clearly not the case in the Ghule and Chavaan (2021) chapter.

5.3 Under-documentation of Sign Languages

Another significant challenge is that the sign languages of the world are under-documented vis-à-vis their spoken language comparators. Thus, even the most inclusive attempts to develop a machine translation product for sign languages is starting from a context where there are significant gaps in the linguistic literature for sign languages (Jantunen et al. 2021). In particular, the complexity of how sign languages use space (e.g. to represent various perspectives on events, and to blend views on events), the use of multiple articulators to present information simultaneously (manual and non-manual simultaneity) (Vermeerbergen et al., 2007), and the potential for partitioning off parts of the body to represent other actors (Dudis, 2004; Leeson & Saeed 2012, 2020), and how these present in a range of registers and/or across text types (Nilsson, 2008) are just some of the issues which are under-described for the sign languages in the SignON project. Jantunen et al. (2021) note that a particular challenge for recognition is handling what sign linguists refer to as 'depiction', which is the visual representation or enactment of an event (Johnston 2012, Fenlon et al. 2014). Depictive content may account for up to 30% of sign tokens "that cannot be unambiguously translated even in the traditional sense" (Jantunen et al. 2021, p. 65).

This under-documentation, coupled with the limited number of deaf/fluent signing members of many machine translation teams working on sign languages, has led to outputs in MT generated content that is very far from the level expected, required, or found acceptable to deaf signing communities (Jantunen et al. 2021).

5.4 The Need for Interdisciplinarity

The siloed nature of initiatives in the field of technological developments for deaf communities is problematic (Bragg et al. 2019; Jantunen et al., 2021). Bragg et al. note that developing successful sign language recognition, generation, and translation systems requires expertise in a wide range of fields, including computer vision, computer graphics, natural language processing, human-computer interaction, linguistics, and Deaf culture (and, we add Deaf studies more generally, along with translation and interpreting studies). However, despite the need for deep interdisciplinary knowledge, they note that "existing research occurs in separate disciplinary silos, and tackles separate portions of the sign language processing pipeline" (p. 16). This prompts three questions:

1) What does an interdisciplinary view of the current landscape reveal?

2) What are the biggest challenges facing the field?

3) What are the calls to action for people working in the field?

To answer these questions, Bragg et al. convened 39 domain specific experts with diverse backgrounds for two days to workshop key issues and present what they suggest are issues often overlooked by computer scientists working on sign languages. They also review the state of the art, identify a set of pressing challenges and outline a call to action for the research community. Given the direct relevance of this work to the SignON project, we take some time to unpack their key concerns here.

The siloed approach to research in sign language processing makes it impossible to respond to the challenges of the field comprehensively (Bragg et al. 2019, Jantunen et al. 2021). Bragg et al. note that, for example, "…there are many computer science publications presenting algorithms for recognizing (and less frequently translating) signed content. The teams creating these algorithms often lack Deaf members with lived experience of the problems the technology could or should solve, and lack knowledge of the linguistic complexities of the language for which their algorithms must account. The algorithms are also often trained on datasets that do not reflect real-world use cases. As a result, such single-disciplinary approaches to sign language processing have limited real-world value (Erard, 2017)." (Bragg et al. 2019, p.16)

SignON seeks to respond to this challenge by ensuring that our use cases have real-world validity, informed by our deaf community partners, stakeholders, and focus group participants. Co-creation principles are embedded into our project and our proposed use cases are outlined in Table 1, below:

1. Every-day communication between a deaf (signer) and a hearing (non-signer) with similar understanding of the conversation setting.

Maria is an industrial engineer. She's the only deaf person in her family, but came in contact with the deaf community when she was a teenager and learned how to sign. Her preferred language is Spanish Sign Language (LSE), and she frequently uses interpreters in both professional and personal settings. Liam has recently moved from Ireland to Spain and started working in the same company Maria does. He never met a deaf person before and knows nothing about SLs. His brief conversations with Maria are difficult. He relies on written text in English which can be impersonal, inefficient and often frustrating for Maria. Liam does not have another option at the moment as he has no time to attend LSE classes because he is learning Spanish (verbal language) at the moment.

The SignON application will provide accessible communication for Maria (as a deaf signer) and Liam (as a hearing non- signer) in every-day communicative situations, where both parties can feel at ease, expressing themselves in their first and preferred language - LSE for Maria and English for Liam.

Benefits: (i) Liam can talk to Maria, and vice versa; (ii) The access to LSE that SignON provides will prime Liam with regard to some of the vocabulary. Limitations and considerations: Position of the mobile device so that Maria can sign using both hands. Mitigation: a small, inexpensive mobile device stand or a case can be used to support the device.

2. Formal dialogues between a deaf (signer) and a hearing (non-signer) with one of them being an expert in the topic of communication; the other is a non-expert.

Joan is a 40 year old woman who uses Sign Language of the Netherlands (NGT) as her primary language with her parents and her deaf friends. Joan has a mild intellectual disability that has made learning to read and write difficult. She has challenges in doing everyday tasks for herself such as lodging money in the bank. When Joan runs into difficulties in these situations, the hearing people she is dealing with often write her notes that she cannot understand. She gets frustrated and regularly gives up on tasks. Regardless, Joan tries to be a fully-fledged member of society and to handle difficult situations, like a visit to the bank, on her own.

In the bank the clerk presents a tablet with **the SignON application installed and personalised** (via SignON's machine learning capabilities) to financial vocabulary (including productive signs). The clerk uses SignON to explain that Joan cannot withdraw more than €750 per day because of the bank's policy. A pre-set help message signs Joan to use the camera to reply. She thanks the clerk and asks if tomorrow she can withdraw more. After the clerk replies she leaves the bank happy she communicated in her native language.

Benefits: (i) Joan understands the regulations and can plan another visit to the bank; the clerk also understands Joan; (ii) Joan is not frustrated. Limitations and considerations: (i) In such situations it is crucial that financial and personal information is conveyed accurately;. (ii) Joan's mild intellectual disability. Mitigation: (i) The clerk should verify Joan's credentials. (ii) The simple interface is easy to use and both Joan and the clerk do not come across issues understanding each other.

3. Facilitating communication between a deaf person using speech and a hearing person.

Russel is a carpenter and he is deaf. He can lipread and talk, although he realises it is not perfect. Sometimes he meets with the clients on-site. Dimitar has recently moved into his new house in Dublin and is replacing the old windows; he has hired the company where Russel works. At the day of installation Russel and Dimitar need to communicate as things have not gone according to plan. For both it is a struggle as Dimitar cannot understand Russel, despite the fact that he is fluent in English; Russel cannot understand (lipread) Dimitar, perhaps due to the fact that English is not Dimitar's first language; Dimitar's knowledge of Irish Sign Language (ISL) is not good enough. Russel has an Android phone with the SignON app installed. He uses **SignON atypical speech recognition** to verbally explain the problem and propose a solution. His message is translated into English (both in audio format and in text) which Dimitar understands. He agrees in English, **which is translated into ISL, signed through the SignON avatar.**

Benefits: Both Dimitar and Russel understand each other and the work proceeds. Limitations and considerations: Inaccurate recognition for either or both atypical and non-native English speech. Mitigation: Russel suggests corrections over the recognised text; Dimitar can type his message.

4. Text to sign language translation for media broadcasting of a crisis situation.

VRT news is the public broadcaster in Flanders, for the Dutch speaking community in Belgium. As a public media company, one of the main goals is to inform society, ensuring inclusion of all its members. A sudden outbreak of a contagious virus in the city centre of Brussels, the capital of Belgium, needs to be reported immediately. Janna (head of the editorial team) directs her team to start a breaking news item on all VRT television channels. She contacts both Efra and Jean; Efra is manager of the Flemish Sign Language (VGT) department, she activates the VGT green key studio and tries to get a VGT signer to the studio as soon as possible; Jean (leader of the subtitling division) activates subtitling workflow. As the subtitle tooling can be activated remotely, a subtitler swiftly acts from a remote location, translating the live speech to text. Getting a VGT signer to the studio at such a short notice is harder as there is a limited group of VGT signers experienced in live broadcast interpreting. It also costs valuable time getting a signer to the VRT premises during a crisis situation. The research and innovation team of VRT has used the open-source framework to provide Efra with a

8. Next Steps

SignON seeks to develop a free, open-source service and framework for conversion between video (capturing and understanding sign language), audio (for speech, including atypical speech) and text, translating between sign and verbal languages, delivered to its users via an easy to use mobile application. To do this, a co-creation methodology is embedded in our process, working with deaf signers and hard of hearing communities to build systems that respond to community needs. To do this, we are currently in the process of securing research ethics permission from Dublin City University and Trinity College Dublin to run our first focus group sessions with participants who are users of Flemish Sign Language, Spanish Sign Language, Irish Sign Language, a wider spread of participants who can communicate via International Sign. Led by the European Union of the Deaf, and our Flemish, Spanish and Irish partners for WP1, these focus groups will allow us to document empirically, for the first time in Europe, the perspectives of deaf signers to machine translation, and what they report as important for community buy-in. This dataset will be coded by SignON researchers and this, along with subsequent rounds of engagement with deaf and hard of hearing participants will allow us to document, analyse, and represent how engaged co-creation methods support processes that facilitate positive outcomes for research approaches in this field. In particular, this approach will feed into the SignON project's recursive engagement approach and guide our project use cases.

---

connector to the SignON Service that will put a virtual signer as an extra video layer on top of the breaking news item on television. The SignON avatar translates the subtitle text into VGT. Valuable information concerning health warnings will now also reach the deaf community, family and friends during the most vital first hours of the pandemic outbreak.

Benefits: Information can be broadcast in a timely manner to all members of society, including deaf society. Limitations and considerations: The quality of signing has to be as good as a human signer. Mitigation: A prerecorded message (in VGT by a human signer) informs the viewers that signing is done automatically and that a human interpreter will be provided as soon as that can be arranged, thus managing their expectations.

9. Conclusions

This report has presented an overview of key challenges that have been documented in the field of machine translation and other technological responses to communities using sign language. In response to these challenges, SignON is committed to a number of principles that reflect state of the art thinking on how to ensure that machine translation projects best serve deaf communities:

- SignON is committed to co-creation, evidenced by the leadership of deaf researchers and deaf community organisations as central to the SignON work, and, through our partners, recursive rounds of engagement with deaf signers from across our target language communities.
- SignON is committed to ethical approaches in our engagement with deaf and hard of hearing communities, with regard to data collection, management, and curation, and in our representation of our research findings.
- SignON is committed to supporting significant developments in the field of machine translation for sign languages that support the real-world use cases which deaf and hard of hearing communities indicate are important to them. Our goal is to work with the communities our project serves, and build towards better societal outcomes, supported by excellent science.

Finally, we have outlined the next steps that we will take in engaging with deaf communities, seeking to explore the current feelings towards machine translation and sign languages in a subset of European deaf communities, and in drawing on these groups to inform our use cases.

References

Albanie, S., Varol, G., Momeni, L., Afouras, T., Chung, J. S., Fox, N., & Zisserman, A. (2020, August). BSL-1K: Scaling up co-articulated sign language recognition using mouthing cues. In *European Conference on Computer Vision* . Springer, Cham. 35-53.

ALLEA (2017) *The European Code of Conduct for Research Integrity*. Berlin: All European Academies.

Bao, P., Maqueda, A.I., del-Blanco, C.R., García, N. (2017).Tiny hand gesture recognition with-out localization via a deep convolutional network. *IEEE Transactions on Consumer Electronics* 63(3), 251-257.

Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., Caselli, N., Huenerfauth, M., Kacorri, H., Veerhoef, T., Vogler, C., & Ringel Morris, M. (2019, 28-30 October 2019). *Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective* ASSETS 19: The 21st International ACM SIGACCESS Conference on Computers and Accessibility, Pittsburgh, PA.

Camgoz, C., Hadfield, S., Koller, O., Ney, H., & Bowden, R. (2018). Neural Sign Language Translation. In *IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT.

Camgoz, Necati Cihan, et al. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Chai, X., Li, G., Lin, Y., Xu, Z., Tang, Y., Chen, X., & Zhou, M. (n.d.). *Sign Language Recognition and Translation with Kinect* IEEE Conference, http://iip.ict.ac.cn/sites/default/files/publication/2013_FG_xjchai_Sign%20Language%20Recognition%20and%20Translation%20with%20Kinect.pdf

Cheng, J., Bian, W., Tao, D.(2013). Locally regularized sliced inverse regression based 3D hand gesture recognition on a dance robot. *Information Sciences 221*, 274-283.

Cui, R., Liu, H., Zhang, C.(2019). A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training. *IEEE Transactions on Multimedia* 21(7), 1880-1891.

Davydov, M. and Lozynska, O. (2017). Information system for translation into Ukrainian sign language on mobile devices. In *2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, Vol. 1. IEEE, 48–51.

De Coster, M., Van Herreweghe, M., & Dambre, J. (2020). Sign language recognition with transformer networks. In 12th International Conference on Language Resources and Evaluation (LREC 2020), Proceedings. Marseille, France: European Language Resources Association (ELRA). 6018–6024.

De Clerck, G., & Pinxten, R. (2012). *Gebarentaal Zegt Alles.* ACCO.

DeMeulder, M., & Haualand, H. (2019). Sign language interpreting services. A quick fix for inclusion? *Translation and interpreting Studies. The Journal of the American Translation and Interpreting Studies Association* https://doi.org/https://doi.org/10.1075/tis.18008.dem de Wit, M., Pérez, S. M., & Peterson, P. R. (2020). *Sign language interpreting on TV and media: sharing best practices. A report of the first European seminar*. www.mayadewit.nl

Dong, C., Leu, M. C., & Yin, Z. (2015). *American Sign Language alphabet recognition using Microsoft Kinect* 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), https://ieeexplore.ieee.org/document/7301347

Dudid, P. (2004). Body partitioning and real-space blends. *Cognitive Linguistics* 15 (2), 223-238.

Elliott, R. ,Glauert,John RW, Kennaway, JR & Marshall, I. (2000). The development of language processing support for the ViSiCAST project. In *ASSETS*, Vol. 2000. 4th.

Erard, M. (2017, 9 November 2017). Why Sign-Language Gloves Don't Help Deaf People. *The Atlantic*. https://www.theatlantic.com/technology/archive/2017/11/why-sign-language-gloves-dont-help-deaf-people/545441/

Fenlon, J., Schembri, A., Rentelis, R.,Vinson, D. & Cormier, K. (2014) Using conversational data to determine lexical frequency in British Sign Language: The influence of text type. *Lingua*. 143, 187-202.

Forster, C. Schmidt, T. Hoyoux, O. Koller, U. Zelle, J. H.Piater, and H. Ney. (2012). RWTH-PHOENIX-Weather: A Large Vocabulary Sign Language Recognition and Translation Corpus. In International Conference on Language Resources and Evaluation (LREC).

Ghule, S., & Chavaan, M. (2021). Implementation of Hand Gesture Recognition System to Aid Deaf-Dumb People. In S. N. Merchant, K. Warhade, & D. Adhikari (Eds.), *Advances in Signal and Data Processing* (Vol. 703, pp. 183-194). Springer. https://link.springer.com/chapter/10.1007/978-981-15-8391-9_14

Grobel, K., & Assan, M. (1997). Isolated sign language recognition using hidden Markov models. IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation.

Gupta, H. P., Chudgar, H. S., Mukherjee, S., Dutta, T., Sharma, K. (2016). A Continuous Hand Gestures Recognition Technique for Human-Machine Interaction Using Accelerometer and Gyroscope Sensors. *IEEE Sensors Journal* 16(16), 6425-6432..

Harris, R., Holmes, H. M., & Mertens, D. M. (2009). Research Ethics in Sign Language Communities. *Sign Language Studies*, *9*, 104-131. https://doi.org/10.1353/sls.0.0011

Hariharan,D. Al-khazraji, S., & Huenerfauth, M. (2018). Evaluation of an English Word Look-Up Tool for Web-Browsing with Sign Language Video for Deaf Readers. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 205–215.

Hill, J. (2020). Do deaf communities actually want sign language gloves? *Nat Electron* (Vol. 3, pp. 512-513). https://doi-org.elib.tcd.ie/10.1038/s41928-020-0451-7

Huenerfauth, M., & Kacorri, H. (2014). Release of Experimental Stimuli and Questions for Evaluating Facial Expressions in Animations of American Sign Language. *Proceedings of the the 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel, The 9th International Conference on Language Resources and Evaluation (LREC 2014), Reykjavik, Iceland.*

Huenerfauth, M., Gale, E., Penly, B., Pillutla, S., Willard,M., & Hariharan, D. (2017). Evaluation of language feedback methods for student videos of american sign language. *ACM Transactions on Accessible Computing (TACCESS)* 10, 1 (2017), 2.

Huenerfauth, M., Zhao, L., Gu, E., & Allbeck, J. (2007). Evaluating American Sign Language Generation Through the Participation of Native ASL Signers. *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '07). ACM*, 211–218. https://doi.org/ http://dx.doi.org/10.1145/1296843.1296879

Jantunen, T., Rousi, R., Raino, P., Turunen, M., Moeen Valipoor, M., & García, N. (2021). Is There Any Hope for Developing Automated Translation Technology for Sign Languages? In Hämäläinen, M., Partanen, N. & Alnajjar, K. (Eds.) Multilingual Facilitation, pp.61-73.

Johnston, T. (2012). Lexical Frequency in Sign Languages. Journal of Deaf Studies and Deaf Education. 17 (2), 63-193.

Kacorri, H., Huenerfauth, M., Ebling, S., Patel, K., Menzies, K., & Willard, M. (2017). Regression Analysis of Demographic and Technology-Experience Factors Influencing Acceptance of Sign Language Animation. *ACM Trans. Access. Comput*, *10*(1), 33 pages. https://doi.org/http://dx.doi.org/10.1145/3046787

Kermit, P., Mjøen, O. M., & Olsen, T. (2014). Safe in the hands of the interpreter? A qualitative study investigating the legal protection of Deaf people facing the criminal justice system in Norway. *Disability Studies Quarterly*. http://dsq-sds.org/article/view/1714/1762.

Kiliboz, N. Ç., Güdükbay, U. (2015). A hand gesture recognition technique for human-computer interaction. *Journal of Visual Communication and Image Representation* 28, 97-104.

Koller, O. , Zargaran, S., Ney, H. & Bowden, R. (2016). Deep Sign: Hybrid CNN-HMM for Continuous Sign Language Recognition. In *British Machine Vision Conference*. York, UK.

Koller, O. , Zargaran, S., Ney, H. & Bowden, R. (2018). Deep Sign: Enabling Robust Statistical Continuous Sign Language Recognition via Hybrid CNN-HMMs. *International Journal of Computer Vision* 126, 12 (Dec. 2018), 1311–1325. DOI: http://dx.doi.org/10.1007/s11263-018-1121-3

Kyle, J. G. (2007). *Sign on Television: Analysis of data based on projects carried out by the Deaf Studies Trust 1993-2005*. https://www.ofcom.org.uk/__data/assets/pdf_file/0015/50181/deafstudies_annex.pdf

Kyle, J., Sutherland, H., & Stockley, S. (2012). *Legal Choices – Silent Process: Engaging legal services when you do not hear*. Research report, Deaf Studies Trust. Available: http://www.deafstudiestrust.org/files/pdf/reports/Legal%20Choices%20Final-sml. pdf

Ladd, P. (2003). *Understanding Deaf Culture: In Search of Deafhood*. Multilingual Matters.

Lang, S., Block, M., & Rojas, R. (2012). Sign language recognition using kinect. *International Conference on Artificial Intelligence and Soft Computing.* 394–402.

Lee, D., Park, Y. (2009). Vision-based remote control system by motion detection and open finger counting. *IEEE Transactions on Consumer Electronics* 55(4), 2308-2313.

Leeson, L. (2019). Ophelia, Emma and the Beast: effortful engaging and the provision of sign language interpreting in emergencies. *Journal of Disaster Management and Prevention*. https://doi.org/ https://doi.org/10.1108/DPM-01-2019-0007

Leeson, L., Flynn, S., Lynch, T., & Sheikh, H. (2020). You Have the Right to Remain Signing: A Snapshot of the Irish Justice System and Deaf Signers. *Teanga*, *Special Issue 11*, 142-173.

Leeson, L. & Saeed, J.I. (2012). *Irish Sign Language*. Edinburgh University Press.

Leeson, L. & Saeed, J.I. (2020) Embodiment in ISL Passives. *Teanga Special Issue* 11, 48-66.

Leeson, L., Sheikh, A. A., Rozanes, I., Grehan, C., & Matthews, P. A. (2014). Critical Care Required: Access to interpreted healthcare in Ireland. In M. Metzger & B. Nicodemus (Eds.), *Investigations in Healthcare Interpreting* (pp. 185-233). Gallaudet University Press.

Leeson, L., & Vermeerbergen, M. (Eds.). (2012). *Working with the Deaf Community: Mental Health, Education and Interpreting*. Interesource Group Publishing.

Liang, Z., Liao, S., Hu, B. (2018). 3D Convolutional Neural Networks for Dynamic Sign Language Recognition. *The Computer Journal* 61(11), 1724-1736.

Mantecón,T.,del Blanco, C.R., Jaureguizar, F., N.García, N. (2019). A real-time gesture recognition system using near-infrared imagery. *PLOS ONE* 14, 1-17.

Mantecón,T., del Blanco, C.R., Jaureguizar, F., García, N. (2016). Hand Gesture Recognition using Infrared Imagery Provided by Leap Motion Controller. In: *International Conference on Advanced Concepts for Intelligent Vision Systems*, ACIVS 2016, pp. 47-57. Lecce, Italy, (2016, October 24-27).

Mathews, E. & O'Donnell, M. (2018). *Reading and Deaf and Hard of Hearing Pupils in Mainstream Education.* Dublin: Catholic Institute for Deaf People, Chime, & the Irish Deaf Society.

Matthes, S., Hanke, T., Regen, A., Storz, J., Worseck, S., Efthimiou, E., Dimou, A.-L., Braffort, A., Glauert, J., & Safar, E. (2012). Dicta-Sign–building a multilingual sign language corpus. 5th Workshop on the Representation and Processing of Sign Languages: Interactions Between Corpus and Lexicon (LREC), Istanbul.

McKee, R. (2014). Breaking news: Sign language interpreters on television during natural disasters. *Interpreting: International Journal of Research and Practice in Interpreting*, *16*(1), 107-130.

Medgadget editors. (2017, 13 July 2017). Low Cost Glove Translates Sign Language, May Be Used to Practice Surgery in Virtual Reality. *Medgadget*. https://www.medgadget.com/2017/07/low-cost-glove-translates-sign-language-may-used-practice-surgery-virtual-reality.html

Molchanov, P., Gupta, S., Kim, K., Kautz, J. (2015). Hand gesture recognition with 3D convolu- tional neural networks. In: *IEEE Conference on Computer Vision and Pattern Recognition Workshops* (CVPRW), pp. 1-7, Boston, MA.

Nan, L., Shourong, W., Minfeng, W., Bo, L., Hong, H., Li, J. (2014). Hand motion recognition based on pressure distribution maps and LS-SVM. In: *International Conference on Mecha-tronics and Control (ICMC)*, pp. 1027-103.

Napier, J., & Barker, R. (2004). Accessing University Education: Perceptions, Preferences, and Expectations for Interpreting by Deaf Students. *Journal of Deaf Studies and Deaf Education*, *9*(2), 228-238.

Napier, J., & Haug, T. (2016). Justisigns: A European overview of sign language interpreting provision in legal settings. *Journal of Law, Social Justice & Global Development: An interdisciplinary journal*, *2*. http://www2.warwick.ac.uk/fac/soc/law/elj/lgd/lgd_issue_2016_2

Napier, J., & Kidd, M. R. (2013). English literacy as a barrier to health care information for deaf people who use Auslan. *Australian Family Physician*, *42*(12), 896-899. https://www.ncbi.nlm.nih.gov/pubmed/24324995

Napier, J., & Leeson, L. (2016). *Sign language in action*. Palgrave Macmillan.

Napier, J., Sabolcec, J., Hodgetts, J., Linder, S., Mundy, G., Turcinov, M., & Warby, L. (2014). Direct, interpreter-mediated or translated? A qualitative study of access to preventive and ongoing healthcare information for Australian Deaf people. In B. Nicodemus & M. Metzger (Eds), *Investigations in healthcare interpreting* (pp. 233–276). Washington, DC: Gallaudet University Press.

Napier, J., & Spencer, D. (2008). Guilty or Not Guilty? An Investigation of Deaf Jurors' Access to Court Proceedings Via Sign Language Interpreting. In D. Russell & S. Hale (Eds.), *Interpreting in Legal Settings* (pp. 72-122). Gallaudet University Press.

Nilsson, A.-L. (2008). *Spatial strategies in descriptive discourse : use of signing space in Swedish Sign Language*. Centre for Deaf Studies, Trinity College Dublin.

Nilsson, A.-L., Turner, G. H., Sheikh, H., & Dean, R. (2013). *An Overview of European Healthcare Provision for Deaf Sign Language Users*.

O'Connor, T. F., Fach, M. E., Miller, R., Root, S. E., Mercier, P. P., & Lipomi, D. J. (2017). The Language of Glove: Wireless gesture decoder with low-power and stretchable hybrid electronics. *PLOS ONE*. https://doi.org/https://doi.org/10.1371/journal.pone.0179766

Ofcom. (2007). *Signing on television: New arrangements for low audience channels*.

Osborne, R., & McPhee, R. (2000, December 12–15, 2000). *Indigenous Terms of Reference (ITR)* 6th UNESCO-ACEID International Conference on Education, Bangkok.

Pabsch, A., & Söderqvist, P. (2015). *UNCRPD Implementation In Europe - A Deaf Perspective: Article 27: Work and employment.* . EUD.

Rose, B. (2020). *Coronavirus: Lack of sign language interpreters leads to legal case against government*.

Science Europe (2018) *Practical Guide to the International Alignment of Research Data Management*. https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

Shi, B., Del Rio, A. M., Keane, J., Michaux, J., Brentari, D., Shakhnarovich, G., & Livescu, K. (2018). American Sign Language Fingerspelling Recognition in the Wild. . *IEEE Spoken Language Technology Workshop (SLT)*, 145–152. https://doi.org/http://dx.doi.org/10.1109/SLT.2018.8639639

Singleton, J., Martin, A., & Morgan, G. (2015). Ethics, Deaf-Friendly Research,and Good Practice When Studying Sign Languages. In E. Orfanidou, B. Woll, & G. Morgan (Eds.), *Research Methods in Sign Language Studies* (pp. 7-20). Wiley & Sons, Inc.

Smeijers, A. S., & Pfau, R. (2009). Towards a Treatment for Treatment: On the Communication between General Practitioners and Their Deaf Patients. *The Sign Language Translator and Interpreter (SLTI)*, *3*(1), 1-14.

Stokoe, W.C. & Wilcox, S. (1995). *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.

Tamura, S., & Kawasaki, S. (1988). Recognition of Sign Language Motion Images. *Pattern Recognition 21*, 343–353. . https://doi.org/ http://dx.doi.org/10.1016/0031-3203(88)90048-9

Tupi, E. (2019). *Sign Language Rights in the Framework of the Council of Europe and its Member States.* Helsinki: Ministry for Foreign Affairs of Finland.

Veale, T., Conway, A., & and Collins, B. (1998). The challenges of cross-modal translation: English-to-Sign-Language translation in the Zardoz system. *Machine Translation* 13, 1 (1998), 81–106.

Verma, H. V., Aggarwal, E., & Chandra, S. (2013). *Gesture recognition using kinect for sign language translation* 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013), https://ieeexplore.ieee.org/xpl/conhome/6695812/proceeding

Vermeerbergen, M., Leeson, L., & Crasborn, O. A. (2007). *Simultaneity in signed languages: form and function*. John Benjamins.

Vermeerbergen, Myriam, Van Herreweghe, M., Smessaert, I. & De Weerdt, D. (2012). "De eenzaamheid blijft": mainstreamed Flemish deaf pupils and wellbeing at school. In Leeson, L. & Vermeerbergen, M. (eds.) *Working with the deaf community: education, mental health and interpreting*. Dublin: Interesource Group Publishing, 101–118.

Wehrmeyer, J. E. (2015). Comprehension of television news signed language interpreters. *Interpreting*, *17*(2), 195-225.

Wheatley, M. & Pabsch, A. (2012). *Sign Language Legislation in the European Union*. Brussels: European Union of the Deaf.

Wilcox, S. E. (2007). *The Gestural Origin of Language*. Oxford: Oxford University Press.

World Association of Sign Language Interpreters (WASLI), & World Federation of The Deaf (WFD). (2015). *WASLI and WFD Guidelines. Communication during natural disasters and other mass emergencies for deaf people who use signed language*. http://wasli.org/wp-content/uploads/2012/11/WFD-and-WASLI-Communication-during-natural-disasters-and-other-mass-emergencies-for-deaf-people-who-use-signed-language-Jan-2015-FINAL.pdf

Young, A., Napier, J. & Oram, R. (2019). The translated deaf self, ontological (in)security and deaf culture. *The Translator* 25 (4), 349-368.

Zafrulla, Z. Brashear, H., Presti, P., Hamilton, H. & Starner, T. (2011). CopyCat: an American sign language game for deaf children. In *Face and Gesture 2011*. IEEE, 647–647.

Zhao, L., Kipper, K., Schuler, W., Vogler, C. Badler, N. & Martha Palmer. (2000). A machine translation system from English to American Sign Language. In *Conference of the Association for Machine Translation in the Americas*. Springer, 54–67.

Zhou, H., Zhou, W., Zhou, Y., & Li, H. (2020, April). Spatial-temporal multi-cue network for continuous sign language recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 07, 13009-13016.

