The EU-Funded Project "Content4all" and Its Options for Accessibility in Broadcast and Educational Settings

Robin Ribback and Michaela Nachtrab
SWISS TXT
Alexander-Schöni-Str. 40
2501 Biel/Bienne, CH
robin.ribback|michaela.nachtrab@swisstxt.ch

Abstract

CONTENT4ALL is a European-funded project with the aim of developing in the long-term a solution which can offer an attractive option for broadcasters to have personalized content for deaf people. The solution offers the possibility of a human sign-interpreter inserted into existing content for the hearing, at a low cost and with no disruption of hearing viewers. The basis for this goal are low-cost remote studios, an inserted 3D model of sign-language interpreters and a database for automatic sign-language translation in structured use cases like weather or sports. This development can also be used in the long-term for the education market, because of similar structured use cases with specific limited vocabulary and technical terms depending on the subjects.

1 Content4all Approach

CONTENT4ALL is a European funded Horizon2020 (H2020-ICT-19-2017) Project with the project No. 762021.

Project partners are Fincons Group AG, Switzerland; University of Surrey, United Kingdom; Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Germany; Human Factor Consulting, Germany; SWISS TXT AG, Switzerland; R.T.I.S.p.A. – Mediaset, Italy

The long-term goal of the CONTENT4ALL approach is to develop a solution which can offer an attractive option for broadcasters to have personalized content for deaf people, where a human sign-interpreter is inserted into existing content for the hearing, at a low cost and with no disruption of hearing viewers.

This approach is based on three phases:

Phase 1: Develop a low-cost remote sign-interpretation and insertion service

This phase includes the development of a) a low-cost remote studio and b) the development of a system to capture and photo-realistically reproduce signs and facial expressions of the human interpreter via a 3D model of the interpreter ("Avatar" of a real person).

Phase 2: Collect, catalogue and analyse sign-language translations in a structured use case

In the short to medium term, data from specific TV programs (e.g. weather, sports etc.) will be collected in order to create a database of subtitles, sign-language manual (hand movement) and non-manual (body gestures and facial expressions) for automatic sign-language translation research.

Phase 3: Automatic sign-interpretation for a structured scenario

In the medium term, there shall be developed an automatic sign-interpretation technology for a limited/structured application scenario, to be developed for more generic scenarios in the long-term.

2 Project description

Improving the accessibility of television content for the Deaf community is an important goal for both EU governments and broadcast industry regulators across the EU. Although legislation is being used to coerce content producers and broadcasters. To do so, the cost of producing sign-language content (both sign-interpreted and sign-presented) and the negative impact of having a sign-interpreter appearing on the content for hearing individuals has relegated sign-language programming to late nights or a small number of sign-presented programs. A low-cost solution for personalized sign-

interpreted content creation can address both problems, leading to greater accessibility to media content for Deaf users.

CONTENT4ALL proposes such a solution to the problem in the short-term, and proposes innovations to technologies that can lead to automated sign-translation capabilities in the long-term. To this end, it builds upon the technologies and expertise demonstrated by the consortium partners in previous EU projects such as ACTION-TV, DictaSign and SCENE.

Here, as a first development, a remote signing apparatus (located off-premises of the television broadcaster) will be produced to capture a human sign-interpreter's signs, pose and facial expressions and to parameterize this information. Afterwards, it will be rendered photo-realistically as a 3D representation of the human sign-interpreter at the broadcaster for the production of the personalized stream for deaf users.

The photo-realistically 3D models have different advantages:

The high quality demanded in the TV sector usually requires very good lighting and a professional studio. This equipment is expensive and the production costs become very high. By using a photorealistic 3D model, instead of a conventional video insertion, these costs can be greatly reduced, Low cost remote studio equipment can be used instead of an expensive TV studio, because the light and the background doesn't have to be always the same.

Sign language is a three-dimensional language. The room, positions of signs and the direction of movement are important elements of sign language. Therefore, a 3D application can help deaf people in better understanding.

The users get familiar to the person who is displayed as the 3D interpreter model. It could have always the same figure and face if desired; even if there are different interpreters behind.

While this solution can be used commercially, the resulting datasets will provide a vast source of information for learning how to parameterize the sign information for translation purposes.

The second development of CONTENT4ALL will focus on advancing the algorithms and models used to do so, with the intent to create an open dataset for further research into automated signing.

SWISS TXT concentrates on the multimodal machine processing of speech using mechanisms of deep learning and AI. The continuous improvement process of the AI models within the framework of an eco-system is integrated into a central workflow system. This Hub records the data during the production process and automatically prepares it for the training of the AI systems. The Training material will automatically have handed over to the rendering engines for the new models.

Finally, CONTENT4ALL will demonstrate automated sign-translation applied to a real-world television broadcasting scenario, which is envisaged to lead to new approaches and innovations in the area in the long-run.

3 Use cases in educational settings

Even if the project itself aims to provide solutions for more accessibility in broadcasting, its results will offer a good basis for the education market.

Education is an area in which improved and comprehensive access for people with disabilities is required more and more.

Deaf students, apprentices and pupils have the right to accessible lectures and demand this, too.

The UN Convention on the Rights of Persons with Disabilities (short CRPD) drives the major market exploitation and corresponding business models for the whole EU & Switzerland.

The CRPD was adopted by the General Assembly of the United Nations on 13 December 2006 in New York. It came into force on May 3, 2008 and today it counts 177 States as Parties, with one peculiarity: it is the first international agreement to which the European Union has acceded. The convention is the first international special convention on the rights of people with disabilities.

The convention was ratified by Switzerland on 15 April 2014 and entered into force on 15 May 2014. By acceding to the Convention, each country committed itself to remove obstacles faced by people with disabilities, to protect them against discrimination and to promote their inclusion and equality in society.

The following articles of the convent are important to the C4A project.

Article 30

Participation in cultural life, recreation, leisure and sport

States Parties recognize the right of persons with disabilities to take part on an equal basis with others in cultural life, and shall take all appropriate measures to ensure that persons with disabilities:

Enjoy access to cultural materials in accessible formats;

Enjoy access to television programs, films, theatre and other cultural activities, in accessible formats;

Enjoy access to places for cultural performances or services, such as theatres, museums, cinemas, libraries and tourism services, and, as far as possible, enjoy access to monuments and sites of national cultural importance.

Article 21

Urging private entities that provide services to the public, including through the Internet, to provide information and services in accessible and usable formats for persons with disabilities; Encouraging the mass media, including providers of information through the Internet, to make their services accessible to persons with disabilities

including providers of information through the Internet, to make their services accessible to persons with disabilities

Article 24

"…States Parties shall ensure an inclusive education system at all levels and lifelong learning" to persons with disability. The state parties have to ensure the equal access to the general education system.

Sign language interpreters on site are currently being used to meet these requirements. However, this support is often organizationally complex and expensive. For this reason, many institutions that must pay for this as well as universities and schools are looking for ways to cost-effectively provide those services.

Technically the advantage of the education market is that there are structured scenarios, too. Each subject at university or school normally contains a limited vocabulary with specific technical terms. Here, too, ap-propriate data can be collected and processed. Via this approach it is possible to start with specific subjects building up the data set for the automatic translation to apply this to more and more subjects in the future.

References

Villa, Matteo et al. 2016. Personalised Content Creation for the Deaf Community in a Connected Digital Single Market (Content4all). Research and Innovations Project Proposal H2020-ICT-19-2017.

United Nations. 2018. Department of economic and social affairs. Convention on the Rights of Persons with Disabilities (CRPD). https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html