



Media Technology Center

Annual Report 2020



“I am very proud of what the MTC has achieved in this short period of time and my gratitude extends to our entire team for their hard work and their excellent accomplishments.”

Prof. Dr. Markus Gross

The digital transformation is profoundly disrupting the way we produce, distribute, and consume media information. Boundaries between media producers and consumers are vanishing in a world where everybody can easily produce media content. We witness the raise of direct-to-consumer digital media distribution platforms as well as the fusion of theatrical and episodic film production. Covid-19 has even accelerated this transformation.

The Media Technology Center (MTC) at ETH was established in 2019 with the generous help of our founding partners from TX Group, Ringier, NZZ, SRG, VSM, viscom, and the ETH Foundation. Our mission is to push the forefront of technological innovation in media to conceive and create new business opportunities as well as to help conditioning traditional business models for the digital world. To accomplish this goal, we have developed a finely tuned process of how we source and run our research projects under the guidance of our stakeholders and academic partners. This model, in combination with a strong commitment to an open innovation policy, allows the MTC to serve as a hub for the Swiss media industry to engage with academia and to envision the future of digital media.

Our steadily growing team of 10 full time equivalent researchers is working closely with technical leaders from industry on topics as diverse as natural language processing, augmented reality, recommendation systems, content analytics, image and video understanding, and much more. Example projects include the Swiss Voice Assistant and Federated Learning.

With the successful hiring of Prof. Ryan Cottrell into the Computer Science Department we completed another important cornerstone of our initial plan: A dedicated faculty for media technology and language processing.

Furthermore, the intersection of media and computer science is resonating extremely well with our students and, besides our well attended seminar series at ETH, we also attracted 17 bachelor and master students. These projects provide a significant leverage of our main funding and allow us to explore promising ideas beyond our main portfolio.

Two years in, we are also at the verge of transferring our technology innovations into the business. In 2020 we released over 40'000 lines of code some of which already being in production testing at our partner companies.

I am very proud of what the MTC has achieved in this short period of time and my gratitude extends to Dr. Severin Klingler and our entire team for their hard work and their excellent accomplishments.

Today, I proudly present you our annual report 2020 and I sincerely hope that you enjoy its reading as much as I do.

With kind regards



Prof. Dr. Markus Gross
Chairman of the Steering Committee

The year in numbers

10

Full time
equivalents

4

Interns



CHF 1M

Annual budget



48'172

Lines of open-
sourced code
published

6

Scientific
publications

7

Events
participated

5

New projects

7

Project videos
released

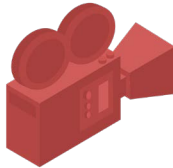
4

Media highlights



521

YouTube views



42

Students
supervised

17

Completed
student projects

Milestones

January | **Isabelle Schrills joins the MTC steering committee, representing Ringier AG.**

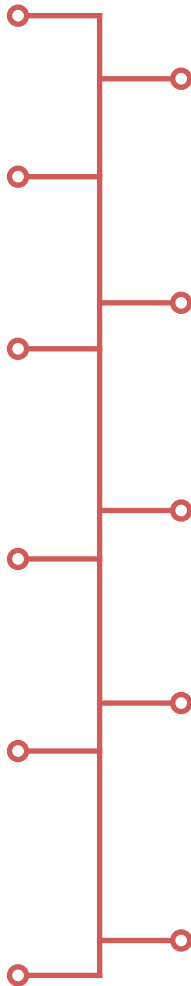
April | **Prof. Ryan Cotterell joins ETH Zurich and the MTC steering committee.**

May | **MTC on YouTube**
The MTC launches a YouTube channel showcasing our work in 2-3 minute animated videos.

August | **Seminar on Media Innovation**
Fully booked, our annual student seminar starts in a hybrid setup (online and physical presence).

October | **Yannick Suter joins the MTC steering committee, representing the TX Group.**

December | **Six scientific publications in 2020**
Throughout the year, researchers at MTC published scientific papers in high-profile venues such as CHI, ICCV, CVPR ICPR, and EuroVis.



January | **First MTC project showcase event**
Open to all our partners, the event features some of the best student projects from 2019 and 2020.

May | **The first top-down projects completed**
The MTC successfully completes the “Hands-free Media” and “Federated Machine Learning” projects.

June | **Five new projects**
The steering committee approves new projects and extends the “Federated Machine Learning” project.

October | **Launch of the MTC Project Hub**
A new web platform that provides easy access to our project portfolio.

November | **MTC receives new funding**
A donation from Huawei allows the MTC to continue to explore key enabling technologies for mixed reality media experiences

People

The Media Technology Center brings together researchers and professionals from the media industry working on applied research projects to investigate future media technologies.

Core Team



Ana Kotarcic



Alberto Pennino



Clara Fernandez



Christian Schneeбели



Elizabeth Salesky



Fábio Porfírio



Graziana Mignone



Julian Mäder



Luca Campanella



Pelin Dogan
Schonbörger



Ribin Chalumattu



Saikishore Kalloori



Severin Klingler



Simone
Schaub-Meyer

Industry Experts



Andreas Häuptli
Verband Schweizer
Medien



Cristina Kadar
NZZ



Gert von Manteuffel
SRF



Isabelle Schrills
Ringier AG



Yannick Suter
TX Group AG

Academia Experts



Prof. Andreas
Krause



Prof. Ce Zhang



Dr. Donald Tillman



Prof. Friedemann
Mattern



Prof. Markus Gross



Prof. Martin Zimmer



Prof. Olga
Sornike-Hornung



Prof. Otmar Hilliges



Prof. Ryan Cotterell



Prof. Thomas
Hofmann

> The MTC's steering committee, composed by the industry and academia experts, is the final authority for all decisions regarding the center.

Building a voice assistant that speaks Swiss German

Millions of users already use voice assistants to get information about the current weather or traffic, listen to the news, play music, and control their smart home. However, none of today's assistants speaks or understands Switzerland's various dialects. In 2020, we built the first voice assistant that can speak different Swiss German dialects.

One of the reasons why there are no smart speakers on the market that understand or speak Swiss German is that Swiss German is a so-called low-resource language, for which only little data and no standardized written form is available. The lack of large data sets makes building text-to-speech models challenging since state-of-the-art models for languages such as English and German rely on big data sets with text and speech aligned.



Our voice assistant prototype in action. Our system can understand and speak in Swiss German. Scan the QR code for a demo video (in Swiss German).



In 2020, the Media Technology Center created the first voice assistant to speak five different Swiss dialects (BE, ZH, GR, VS and SG) using state-of-the-art technologies in neural speech synthesis. We explored methods that learn how to pronounce Swiss German for all different dialects at once, thus potentially reducing the amount of data needed to train these synthesis models. Our models mimic real speakers' voices and are trained on parallel text (audio recordings and text pairs) of about 3 hours per dialect.

Thanks to a collaboration with recapp AG that provides speech recognition services for Swiss German, our voice assistant can understand Swiss German as well. The jointly developed prototype answers everyday requests such as reading the news headlines and weather forecasts in any of the five dialects currently available.

Many applications beyond voice assistants

At the core of this project, we built a text-to-speech model for different Swiss German dialects. Our model first translates High German



text into Swiss German text using machine translation methods. The Swiss German text is then converted into Swiss German speech using a neural speech synthesis method. In addition to the voice assistant prototype, we deployed a Swiss Voice REST API that takes High German as Input and produces Audio in all supported dialects. Our easy-to-use API allows you to apply our technology stack for many different contexts such as article narration, podcast generation, or the generation of individualized radio moderations.

Large public data set

As part of this project, we collected over 3 hours of voice recordings per dialect together with Swiss German and High German transcripts. To foster further research, we will release this data set, making it the most extensive corpus of parallel speech-text data for Swiss German (in terms of the number of sentences). The data set will be freely available to the research community. Researchers can

use our data set to explore speech-synthesis methods and other technologies dealing with Swiss German texts, such as dialect identification, machine translation, or even to explore linguistic properties of the various dialects. ■

- › Our system can generate speech in five different Swiss German dialects based on High German input text with three more dialects (AG, BS, LU) available in Q1 2021.
- › In collaboration with recapp AG, we created the first voice assistant to understand and talk different Swiss German dialects.
- › We created a high-quality data set with over 30 hours of speech samples from different Swiss German speakers.
- › We released a web service that can translate High German text to Swiss German speech automatically.

Technological Foundations of Hands-Free Media

Smart glasses and augmented reality (AR) headsets provide the opportunity for radically new ways of media consumption. In 2020, we developed different prototypes for potential future hands-free media experiences to explore the potential and address the challenges of this new form of media consumption. Our most ambitious project was to build a streamlined pipeline to generate city-wide AR content for the city of Zurich.

Embedding media into the real world allows for new immersive experiences that overlay information in a context- and location-aware way and provide a rich hands-free media experience. While the opportunities are vast, many open technological challenges remain. Only after overcoming these challenges it will become clear which forms of hands-free media consumption are possible.

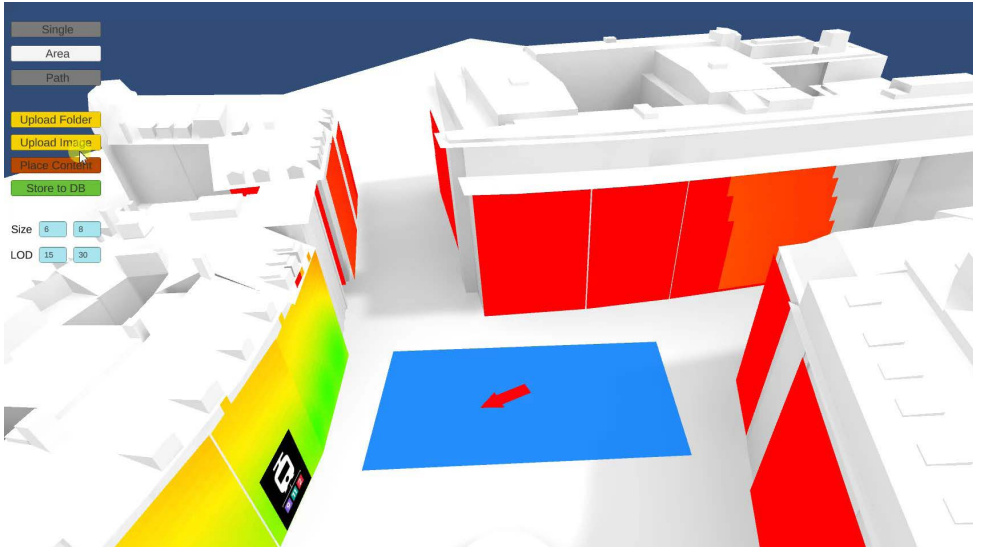
The experience with augmented reality (AR) applications hinges on the availability of exciting content to consume. At the moment, relatively few compelling AR experiences exist because it is still very challenging to create gripping AR content, especially if the goal is a city-wide experience. Generating an AR experience typically requires good programming and 3D modeling skills, which severely limits who can create AR content. In this project, we wanted to address this challenge by building a new pipeline that allows novice AR users to generate and prototype compelling AR content quickly.

Content creation made easy

We developed a very streamlined way to generate AR content in the city of Zurich. Using a VR device, the content creator is virtually placed on a square in the city of Zurich. The user can then freely place 2D and 3D content and draw 3D shapes. All creations are automatically transferred to the correct locations within Zürich and can be experienced by other users using their smartphone or compatible Augmented Reality headsets. We believe that this allows artists and content creators to conceive impressive artistic installations within Zürich at ease. Also, it allows for fast prototyping of AR applications and ideas that require no programming knowledge or prior experience with AR technology.

Automatic placement of virtual billboards

We developed a tool that can automatically place virtual 2D billboards within the city of Zurich. The system automatically projects the 2D content at building walls to blend the real-world environment with the virtual content



[top] Our web tool automatically places 2D content on walls of buildings for best visibility. We compute a suitability score for each content ranging from red (location is not suited for content) to green (location is very well suited for content).

[right] The new augmented reality WYSIWYG (what you see is what you get) pipeline. Augmented reality content can be easily placed in virtual reality.





3D and 2D content that was created by the user is automatically placed in AR within the city of Zurich and can be visualized using a smartphone or a Microsoft HoloLens headset.

seamlessly. Our system automatically sets the virtual billboards in places that pedestrians easily see and optimizes it for other factors, such as distance to the viewer. Our application renders content at three different levels of detail. From a distance, it provides just the most essential information and as the user gets closer to the content, it provides a detailed view. For example, the user would only see the headline from an article when far away from a billboard and moving closer to the billboard the user would see a teaser image and a teaser text.

Automatic placement in the city of Zurich

Our framework automatically places any content created with the methods described above within the city of Zurich. Currently, users can test the experience either using a smartphone or the Microsoft HoloLens. The exact user location is automatically computed based

on so-called spatial anchors that allow us to identify the user location and put all content at the correct place. ■

- › Researchers at MTC built an easy-to-use infrastructure to quickly and easily prototype AR applications without requiring coding or any previous AR experience.
- › Our web tool allows for the automatic placement of virtual billboards in the city of Zurich. Content visualization is automatically optimized to align with buildings in the town to make for a seamless viewing experience.

Is Federated Machine Learning a solution to overcome data silos?

While data is abundant, it is typically stored in data silos and cannot be shared between media houses. How can we overcome such data silos? In this project, we explored a new technology called Federated Machine Learning as a potential solution.

Federated Machine learning allows companies to collaboratively build machine learning models without the need to share data at any point. A model trained on many different data sets should be more robust and provide better accuracy for tasks such as predicting user conversion or providing article recommendations. This project explored how to leverage federated machine learning for building shared models for the media industry. For this, we developed a prototype system that allows the MTC and our partners to train machine learning models collaboratively. The framework is available for free as an open-source project for commercial and non-commercial use. The system has been deployed and tested with two of our partners TX Group and NZZ.

Together with our partners, we investigated the benefits of a federated machine learning setup for two different use cases. First, we explored the benefits of federated learning to predict user conversions (user subscribing to the website). Second, new models for recommending news articles to users based on a federated learning setup were developed and tested. As part of these efforts, we explored how to quantify the improvements from indi-

vidual data sets and define a fair cost-sharing model between the partners.

Our findings show that it is technically feasible to set up a federated learning framework that allows for secure and private model building jointly with different stakeholders. Federated model performance compared to individual model performance improved by up to 20% in our experiments. Depending on the use case, 20% might provide a significant return on investment and make a federated learning setup an appealing choice. ■

- › Our prototype Federated Learning framework is successfully running on the infrastructure of NZZ and TX Group.
- › Performance of federated models improved by up to 20% compared to models trained on data from a single company.
- › The framework has been released as open-source and is available for commercial and non-commercial usage.



Education

Besides our research projects, we are committed to introduce students to the vast possibilities in media technology and benefit from fresh ideas students are bringing.

Seminar on Media Innovation

In 2020, 24 students participated in our media technology seminar. We introduced the students to the latest research in different areas of media technology. It is an exciting field laying at the intersection of computer vision, computer graphics, natural language processing, and machine learning. The seminar covered a broad spectrum of topics considering the technical innovations and the possibilities these technologies provide to professionals in the media industry and consumers of media.

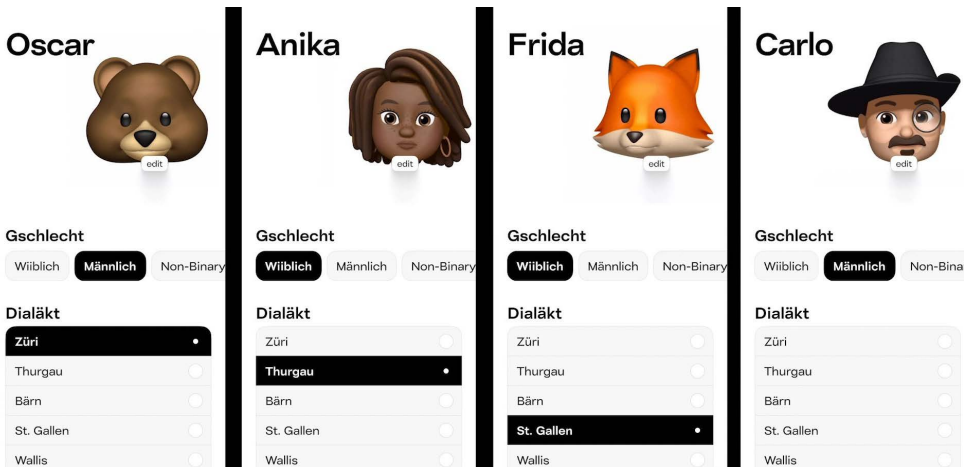
Each student presented a research paper on topics ranging from recommender systems, automatic text generation, video captioning to advances for intelligent augmented reality interfaces.

Bachelor and Master theses

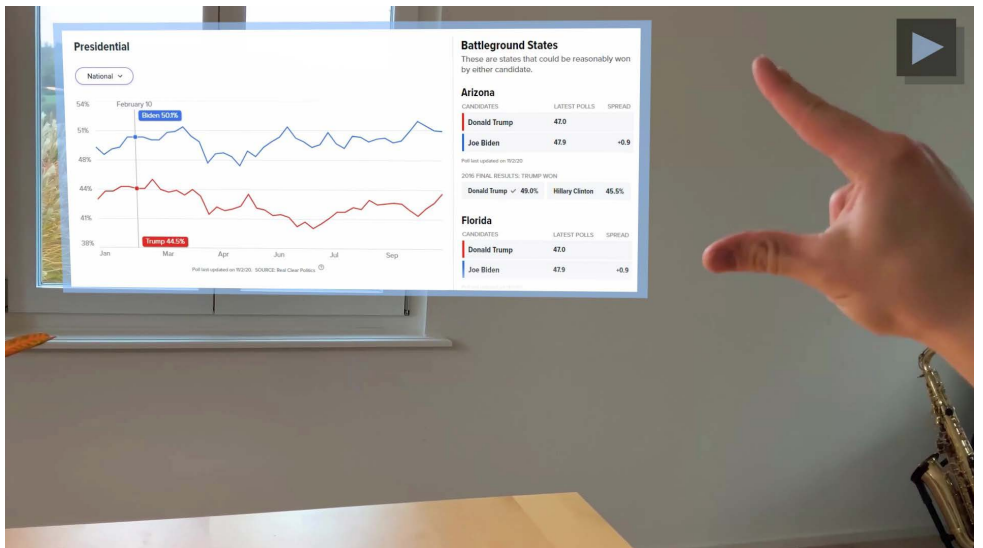
Our researchers supervised a total of 17 Bachelor and Master theses in 2020. In these hands-on projects, students are programming new algorithms, experimenting with state-of-the-art code, and conducting experiments to validate their ideas. The projects conclude with a written report providing all the details of the students' approaches and the experiments they conducted. You can find some of the highlights of these projects on page 16.

Design workshops at the Zurich University of the Arts

To allow for better exchange between technology and design. The Zurich University of the Arts invited the MTC to participate in two design workshops. The goal of the workshops was to design product ideas based on the technologies developed at the MTC. Over 15 students from different study directions in the area of design participated in these workshops. Product prototypes were pitched and demonstrated in a semi-public event open to our partners and members of ETH and ZHdK. ■



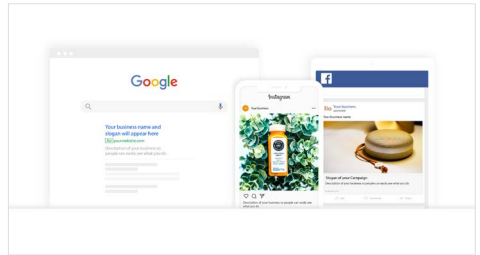
SAEPP. Everyone has different needs when using a voice assistant. The ZHdK students Tatjana Streit, Lia Studerus, Lucie Bachmann, and Fabian Frey developed a concept for an entirely personal voice assistant experience powered by the MTC Swiss Voice technology.



Seamless Media. Augmented reality provides all new possibilities for media consumption. Vera Gut, Dominic Platten, Sikharin Meier from ZHdK created a concept for seamless media consumption in an augmented reality space. Switch seamlessly between different media formats (audio, video, text) while continuing in the story exactly where you left off.

Bottom-up innovation

Innovation often starts small. In these bottom-up projects, we explore new technologies together with Bachelor and Master students. Seventeen students completed their projects at the MTC in 2020. This year was especially challenging for our students, without access to our lab and remote supervision only. We are very proud of what our students were able to achieve within 3-6 months.



Argument Browser - A Tool for Pluralistic Discourse

by Luca Campanella, Dr. Severin Klingler

Interesting arguments are easily buried in heated conversations in comment sections. We created a tool that allows users to discover interesting comments that enrich the conversation.

Automatic Stylised Text Generation

by Nicholas Dykeman, Dr. Severin Klingler, Leon Palaić

Automatic text synthesis for a given text style such as to inform, convince, tell a story etc, by extracting text styles from product reviews and online advertisements.

Learn about the remaining theses on our website.

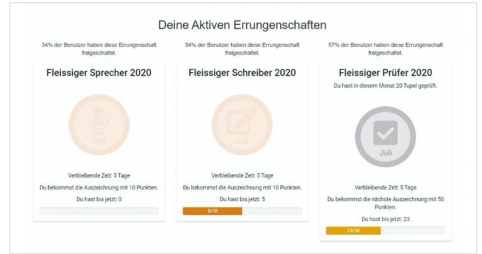




Code-Mixed Speech Synthesis

by Jonas Stehli, Dr. Pelin Dogan Schonb rger

Exploring a neural framework for a consistent and natural synthesised speech for English words in Swiss German text.



Data Collection for Swiss Voice Assistant

by Viturin Z st, Dr. Severin Klingler

Developing a gamified platform for the crowd-sourced collection of Swiss German speech and text.



Detection and Visualization of Safety Critical Objects in Panoramic Images

by V ronique Kaufmann, Prof. Otmar Hilliges, Dr. Simone Schaub-Meyer, Dr. David Lindbauer

A simulation and evaluation tool for detecting moving objects in a 360  video stream, which can then be used to visualize the current situation and to detect possible collisions.



News Globe

by Nicholas Ingulfsen, Dr. Simone Schaub-Meyer, Dr. Tobias G nther

Creation of a virtual globe that shows what is currently happening in the world, providing new ways to interact with news stories and visualize relations between events in an intuitive and informative way. ■

Outreach

Due to the global pandemic not all outreach activities planned for 2020 could take place. The MTC quickly moved focus to online channels and participated in 6 virtual events, started a YouTube channel and newsletter.

Youtube Channel

In June 2020, the MTC started a new YouTube channel. The channel publishes bite-sized videos about finished projects and highlights exciting work in progress. A total of 7 videos were published in 2020.

The MTC on YouTube:



Events

[23 Jan] **MTC Project Showcase @ Presseclub (TX Group)** – Event open to our partners with short project pitches from our bottom-up projects and live demos for the audience to explore.

[29 Jan] **MTC Highlights @ Kanti Wolfbach (Jung von Matt)** – Presenting Highlights from the first year of MTC to the staff of Jung von Matt.

[28 Apr] **MTC Highlights @ Migros Markom „Media Technology & Innovation“** – Presenting Highlights from the first year of MTC to the staff of Migros Markom.

[22 Sep] **Case Study: SRF - Swiss Voice Assistant @ ETH Industry eWeek 2020** – Presenting the work and collaboration on the Swiss Voice Assistant project, alongside Dr. Christian Vogt (SRF).

[1 Oct] **MTC Highlights @ HackHackers Meetup in Zürich** – Introducing MTC and projects to the Swiss HackHackers crowd consisting of hacks (journalists) and hackers (computer scientists).

[23 Oct] **Panel of Media Labs @ Zukunftswerkstatt 2020, University of Zürich** – Panel discussion on media labs as a new form of an innovative work environment.

[5 Nov] **“ETH Media Tech Center – A place of open innovation in the Swiss Media Industry” @ TX | Conference** – Presenting an overview of the work at MTC and opportunities for collaboration as part of the virtual TX | Conference.

Scientific Publications

Enriching Video Captions With Contextual Text *P. Rimle, P. Dogan, M. Gross*
International Conference on Pattern Recognition 2020 (ICPR)

Image Reconstruction of Tablet Front Camera Recordings in Educational Settings *R. Wampfler, A. Emch, B. Solenthaler, M. Gross*
2020 International Conference on Educational Data Mining (EDM)

Simplifying the Process of Creating Augmented Outdoor Scenes *R. Chalumattu, S. Schaub-Meyer, R. Wiethuchter, S. Klingler, M. Gross*
2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)

Glyph-Based Visualization of Affective States *N. Kovacevic, R. Wampfler, B. Solenthaler, M. Gross, T. Günther*
EuroVis 2020

Affective State Prediction Based on Semi-Supervised Learning from Smartphone Touch Data *R. Wampfler, S. Klingler, B. Solenthaler, V. R. Schinazi, M. Gross*
SIGCHI Conference on Human Factors in Computing Systems (CHI 2020)

Style Transfer for Keypoint Matching under Adverse Conditions *A. Uzpak, A. Djelouah, S. Schaub-Meyer*
International Virtual Conference on 3D Visiton (3DV 2020)

MTC in the news

[26 Mar] **Machines learn to speak Swiss German** – ETH Industry News Video

[14 May] **The virtual made real – new technology for the media of the future** – News article, News portal ETH Zürich

[12 Oct] **Story on our “Augmented Paper” project** – TV coverage on BBC4 Tech (Aired on BBC World News network in Arabic speaking countries)

[2 Dec] **Parler les dialectes alémaniques grâce à l’IA** – Radio coverage on RTS of our Swiss Voice Assistant project

Technology transfers: moving from technology to products

True to our mission to shape the future of media technology, we are committed to open innovation and to enable technology transfers from academia to industry.

In 2020, the first two top-down projects and many bottom-up projects were concluded successfully. Based on the developed research prototypes, we started several initiatives to explore commercialization and technology transfer possibilities with our partners.

MTC Project Hub

To make our project portfolio visible to all our partners, we created the MTC Project Hub at <https://projects.mtc.ethz.ch>. The new web platform provides an overview of all our projects and serves as the first step for successful technology transfer by providing detailed project information, including demo material such as videos and presentations. Our partners are currently reviewing our portfolio for technology transfer opportunities.

Commitment to open source

At the MTC, we strongly believe in open innovation. Therefore we make our code and software available under open source licenses to the public. Our software is available for commercial and non-commercial purposes at no additional cost. In 2020, we already released three projects as open-source, with more to come in 2021. With our commitment to open-

source, we established the first step for successful technology transfers from academia to industry. For the Federated Learning project, code from the MTC is already used by our partners today.

Joint proof-of-concept as a way to enable tech transfers

We believe that fast prototyping cycles to build initial proofs-of-concept can go a long way in making new technologies more accessible. We collaborated with the startup company recapp AG for the Swiss voice project to create a Swiss Voice Assistant prototype that can understand and talk Swiss German. This jointly developed proof-of-concept demonstrated the feasibility of developing a Swiss voice assistant. Using the proof-of-concept, we are currently exploring possibilities to commercialize the underlying technologies. ■



Project Overview

Here you find direct access to all the projects.



Technological Foundations of Hands-Free Media
by Dr. Simone Schaub-Meyer, Ribin Chalumattu

Augmented Paper
by Yannick Roskopf, Dr. Simone Schaub-Meyer

Speech Synthesis Models for Germanic Low-Resource Languages
by Dr. Pelin Dogan, Julian Mäder, Philippe Goetschmann, Agon Serifi, Philipp Rimle

Knowledge Aware Web Content
by Robin Bisping, Dr. Severin Klingler

Argument Browser - A Tool for Pluralistic discourses
by Luca Campanella, Dr. Severin Klingler

Automatic Stylised Text Generation
by Nicholas Dykeman, Dr. Severin Klingler, Leon Palaić

Contextual Video Captioning
by Philipp Rimle, Dr. Pelin Dogan

A Practical Federated Learning Framework for Small Numbers of Stakeholders
by Dr. Saikishore Kallioor

Virtual News Globe

The MTC project hub is an online platform accessible to all our partners providing an overview of our project portfolio. The project platform was established as a support tool for our technology transfer efforts and to easily provide information, documentation and links to available code for our partners. The project platform serves as the basis for technology transfer discussions.

ETH Zürich
Media Technology Center
STD
Stampfenbachstrasse 48
8092 Zürich

mtc@ethz.ch
www.mtc.ethz.ch

The Media Technology Center is supported by Ringier, TX Group, SRG SSR, NZZ, Verband Schweizer Medien, viscom and the ETH Zurich Foundation.