



# Media Technology Center

Annual Report 2021



“I am proud and grateful to accompany MTC’s pathway as a member of its Academic Steering Committee. The early years were successful. Let’s ignite the next stage of the rocket.”

Prof. Dr. Martin Zimmer

Never before the world had access to so many stories at the same time, regardless of where they are happening. Media technologies have changed and evolved. But what touches, excites, activates and entertains people has not changed at all: we are still the same human beings. Artists, designers, writers and musicians know that.

When informatics team up with them, media innovation is born. This is exactly the path that the MTC has taken in 2021. A signed research collaboration between Zurich University of the Arts and ETH Zurich made this possible and is a milestone for the vision of the MTC: being an innovation platform at the intersection of technology, design and journalism.

Talented arts and design students were connected with internships or diploma theses with the center. In an eight-month collaboration, the augmented reality project «Talking Houses» was created. It was stimulating to watch two computer scientists, an animation artist, a musician and a designer work together on a prototype. They first had to agree to a common language. As soon as it was found, they inspired each other.

The MTC's 2021 annual report is impressive. The center's efforts around Human-centric AI increased in 2021. Making complex machine learning pipelines easily accessible through web technologies is its current focus.

In addition, AI offers great opportunities in the field of natural language processing (NLP). With Prof. Ryan Cotterell, the MTC has an international NLP-expert in its Academic Steering Committee. Text is at the heart of what our

media partners are doing – and NLP will accompany us in the years to come.

As MTC we are committed to introduce a young generation to the challenges in media technology. In 2021, twelve students worked on their own projects and seven internship positions have been filled with young talent. MTC hosted a seminar with twenty-three students on latest research in media technology and its applied practice in the industry as well as a "Media Innovation Lab" with eighteen design students at Zurich University of the Arts.

I am proud and grateful to accompany MTC's pathway as a member of its Academic Steering Committee. The early years were successful. Let's ignite the next stage of the rocket.

I would like to thank Prof. Dr. Markus Gross for his time and energy as Chairman and Dr. Severin Klingler for his work as Managing Director. Congratulations to the whole team for their success, working competently and intensively, but also having a lot of fun in 2021 when the Covid-19 situation allowed it.

Kudos to you!

With kind regards

Prof. Dr. Martin Zimper  
Academic Member of the Steering Committee

# The year in numbers

13.5

Full time  
equivalents

7

Interns



CHF 1.2M

Annual budget



9'000

Lines of open-  
sourced code  
published

8

Scientific  
publications

8

Events  
participated

12

Students  
supervised

4

New projects

36

Released hours  
of Swiss German  
speech data



# Milestones

January | **MTC's first doctoral student**

Eleftheria Tsipidi starts her work on multilingual summarization supervised by Prof. Ryan Cotterell.

April | **SwissDial dataset**

We released the largest parallel corpus of Swiss-German text and Swiss-German speech, containing three hours of high-quality audio recordings for eight dialects.

May | **New offices**

The MTC and the Game Technology Center (GTC) move into a new ETH building, promoting the synergies between the two centers.

September | **MTC @ Applied Machine Learning Days (AMD L 21)**

Presentation on how the MTC together with its industry partners is pushing boundaries for media technology.

November | **New projects approved**

Our steering committee approves three projects in the areas of text summarization, speech synthesis and text classification for German.

March | **Release of two data annotation tools**

The tools for image and text annotations were released with the research community under an open-source license.

April | **MTC receives new funding for a 3-year InnoSuisse project**

With a focus on intelligent AR characters, the project is executed with the Game Technology Center (GTC), Disney Research Studios and Korean universities.

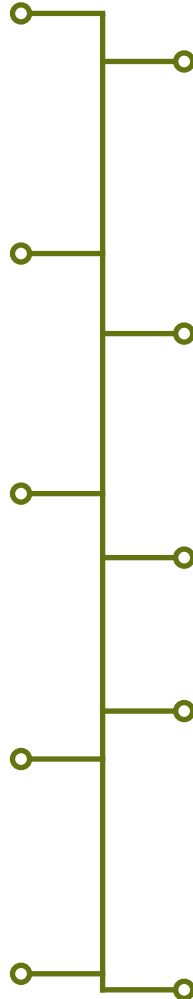
July | **Swiss Voice project wins the Innovation Award at the SwissText Conference 2021**

October | **Video synthesis work at the Digital Festival Zürich**

Our AI tool for generating synthetic videos of SRF moderators is featured at the SRF booth.

December | **Eight scientific publications in 2021**

Throughout the year, researchers at MTC published scientific papers in venues such as 3DV, ASRU, FEVER, RecSys, and WSDM.



# People

## Core Team



Alberto Pennino



Ali Uzpack



Ayça Takmaz



Clara Fernandez



Christian Schneebeli



Daniel Vera Nieto



Eleftheria Tspidi



Fábio Porfírio



Julian Mäder



Laura Mascarell



Marc Willhaus



Philippe Schlattner



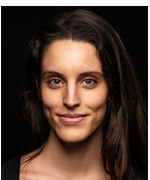
Ribin Chalumattu



Saikishore Kalloori



Severin Klingler



Silvia Monti



Tatyana Ruzsics



Thomas Steinmann

## Industry Experts



Cristina Kadar  
NZZ



Gert von Manteuffel  
SRF



Isabelle Schrills  
Ringier AG



Stefan Wabel  
Verband Schweizer  
Medien



Yannick Suter  
TX Group AG

## Academia Experts



Prof. Andreas  
Krause



Prof. Ce Zhang



Dr. Donald Tillman



Prof. Markus Gross



Prof. Martin Zimmer



Prof. Olga  
Sornike-Hornung



Prof. Otmar Hilliges



Prof. Ryan Cotterell



Prof. Thomas  
Hofmann

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

# AI Meets Design

The MTC was conceived from the beginning as an innovation platform at the intersection of technology, design, and journalism. We are very proud that in 2021 we were able to move one step further with our vision by establishing a close collaboration with the Zürich University of the Arts.

Bringing different disciplines together to work jointly on future media technologies and new media experiences is an essential part of MTC's vision. We are, therefore, delighted that in 2021 we were able to set up a research collaboration agreement with the Design Department of the Zurich University of the Arts (ZHdK). This collaboration, led by Prof. Dr. Martin Zimmer (Cast / Audiovisual Media) on the side of ZHdK, allows the MTC and ZHdK to work more closely together. It will enable us to bring in people from the Design Department of ZHdK to work together with our AI researchers on various projects. In the following, we show some of the first highlights of this collaboration in 2021.

## Joint research projects

ZHdK is supporting the research projects at MTC with its expertise in storytelling, video, animation, interaction design and user experience. As an essential part of these joint projects, we are bringing in designers and artists from ZHdK into the center to work in close exchange with our researchers. As far as COVID rules allow, artists and designers spend at least half a day per week physically in the lab of MTC to establish a mutually beneficial feedback loop between lab artists and AI researchers. We believe that the collaboration

at the intersection of AI and arts will create thrilling ideas. In 2021, we started the first joint project with the animation artist François Chalet, who worked closely with the augmented reality (AR) team at MTC to develop the AR experience "Talking Houses" in the old town of Zurich (Niederdorf). Users will be taken on a guided walk through the old town. Interesting buildings will be brought to life as animated characters and start talking to the user to tell their stories from the past and present. The AR experience includes a scavenger hunt. It is tightly connected with the overall AR projects at MTC to build towards a vision of a seamless city-wide AR experience.



First impression of the guided augmented reality walk in which buildings are brought to life through animated characters.



### Media Innovation Lab

For the third time, MTC was invited to participate in the bachelor module «Media Innovation Lab» at ZHdK. The idea of the lab, led by Prof. Martin Zimper and Creative Director Alex Hefter, is that students, together with the support from researchers at MTC (Ribin Chalumattu, Philippe Schlattner, Ayça Takmaz), create media content and media formats based on technologies developed at the MTC. Students were encouraged to push boundaries, and experimentation was explicitly encouraged. Since students from all design fields participated, they brought in diverse design skills and artistic ideas that led to very creative project outcomes. This year eighteen students explored new AR experiences based on our work on city-wide AR and intelligent AR characters. Concepts ranged from a city-wide horror game experience (Kathleen Boren, Florentin Erb, Michael Kämpfer, Frederik Vestergaard) to virtual pets for children (Aylin Acikel, Kimon Apeltsotou, Sasha Luo, Michelle Hunziker, Manu Wirth). Another group focused on how a creative tool could use our technology for emotion and stance prediction to positively and humorously interact with the often hateful online discourse (Sebastian Burckhardt, Nemo Brigatti, Sibel Metinoglu, Katharina Sauter), and “Pigeon” created an AR indoor experience for exhibitions or project presentations (Lucienne Chrétin, Ryan Bran, Felix Brunhold, Pietro Peduzzi, Andreas Waldburger). Product prototypes were pitched and demonstrated in a semi-public event at the end of the lab.



In **Rayline**, an immersive AR experience, users explore an alleyway in the old town of Zürich that is transformed into a spooky, dimension-warping experience.



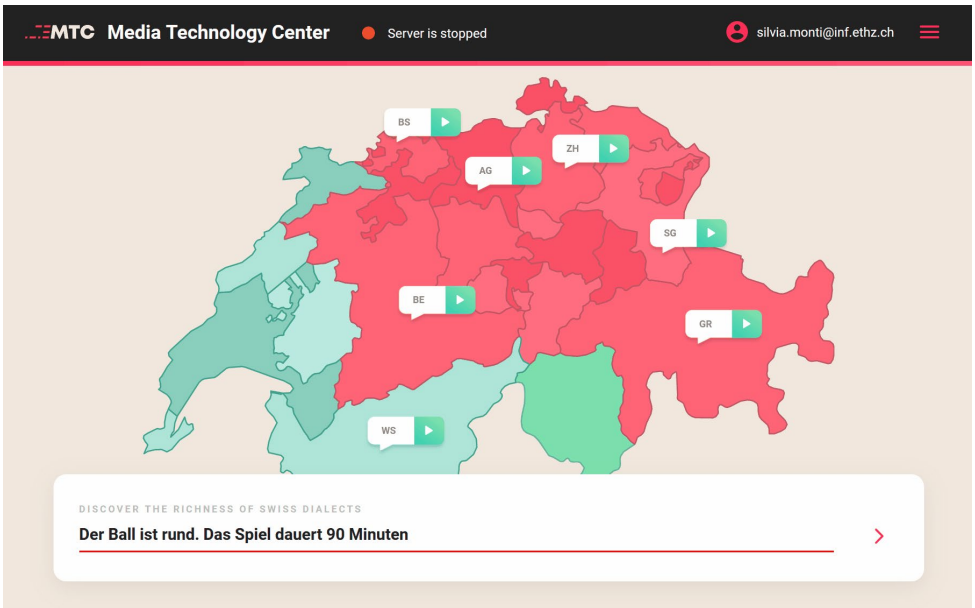
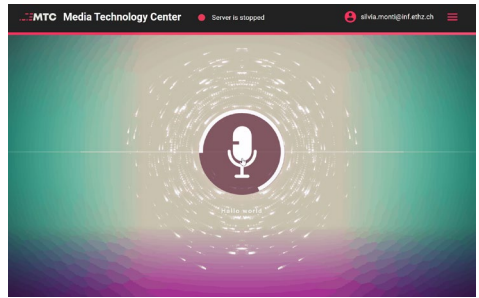
In **Mago**, plushies are turned into playful animal companions for kids. With the help of AR and an understanding of their surroundings these virtual pets can interact, react and play, but also tell children when it is time to go to bed.



**Niceomator** is a twitter bot that uses natural language understanding algorithms to automatically detect hurtful tweets, and answers in a variety of ways: either with a meme, rewording the original tweet or providing a suggestion. The example shows the twitter bot suggesting that a user meditates.

## Design interns at MTC

Based on past media innovation labs and joint workshops, we are inviting ZHdK students with exciting ideas to join MTC for an internship to bring the ideas to life with our AI team. These internships have a duration of three to six months. An important goal for us is that the designer or artist is working in close collaboration with our technical team to exchange ideas and develop novel solutions or use cases for the technology. In 2021, we had the pleasure to work with Silvia Monti, who joined us for six months from June-December. She worked on new ideas and user interfaces for our Swiss German speech synthesis project.



Inspired by an installation in the Museum for Communication in Bern, Silvia Monti (ZHdK graduate) developed, as part of her internship at MTC, a web tool that interfaces with our Swiss German speech synthesis technology in new ways and allows you to beautifully explore differences in Swiss dialects.

### Joint student thesis

We are encouraging and supporting students from ZHdK and ETH to work on diploma projects and theses together. In 2021 Paulina Zybinska from ZHdK and Alessia Paccagnella from ETH finished their joint project to create an exhibition using video retargeting. The goal was to create an interactive experience that confronts visitors with videos showing themselves expressing opinions and beliefs that visitors would not express themselves. The setup consists of a first phase, where the visitor answers a couple of questions while being recorded. In the second phase, the visitor enters a room that shows big screens (mimicking mirrors) that show deepfake videos of the visitor saying things they never said. The work has been presented under the name “Faketual Reality” during the MA Interaction Design Exhibition at ZHdK during summer 2021. ■



**Faketual Reality** The audiovisual installation uses video and audio-controlled synthetic faces to give the visitors a disturbing experience of facing their alter egos. Based on a «Big 5 Traits» personality test and on-site recordings, simulated realities are presented to the visitors as short video clips with synthesized speech.

# Individualized video moderations

Today, algorithms automatically produce data-driven news articles for finance, sports, elections, or national votes. These algorithms can deliver individualized and hyper-local information to the reader. Recent advances in video synthesis technologies may allow us to automatically move from generated text to generated video moderations. We built a prototype system that produces personalized moderations based on audio or text.

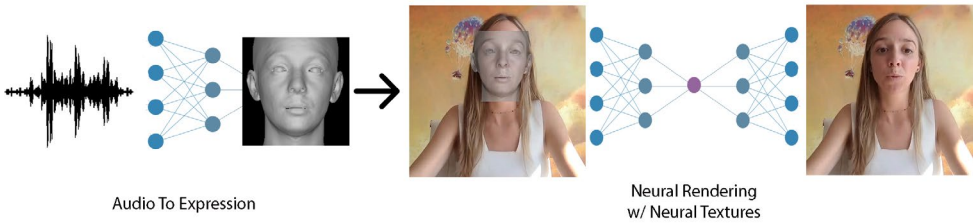
Text generation algorithms can produce text for small and specific events that journalists typically do not have the time to write about. These technologies also allow contextualizing news to your background: from niche sports events in your neighborhood to election results for your municipality. For example, an article about election results might compare and contrast the nationwide results to the results in your municipality. Text generation technologies, therefore, allow for individualized and hyper-local news content. The Washington Post used such technologies to generate audio updates on the 2020 US election automatically, and Amazon built a system that automatically produces sports narratives for commentators to use during sports games.

Recent advances in video synthesis have the potential to take this one step further and create individualized video content. These new methods can create videos of a person speaking new content that was never recorded. To achieve this, a real person's head and lip movement is learned based on short video clips of that person. New movements can then be synthesized given a new text or speech. These

technologies have tremendous potential for creating individualized moderations (e.g., your personalized news anchor, weather anchor), updating moderations faster and at a lower cost or injecting standard news formats with local information at scale. In this project, we explore the technology to generate synthetic weather moderations and announcements of election results and build a prototype system for personalized moderations.

## **High-definition lip synchronization**

Research on video synthesis methods typically focuses on relatively low resolution output. In 2021, we adapted state-of-the-art lip synchronization technologies to produce high resolution outputs. Given that the lip movements are strongly correlated to the speech, our system can learn such a mapping from speech to lip movement. Our pipeline employs a 3D face model that is used to map the input audio to deformations of the 3D model. This mapping can represent person-specific talking styles. In a second step, a neural rendering network generates photo-realistic video content reproducing the person-specific appearance and moving the lips according to the



Given a speech recording our pipeline first generates a 3D model of the head with the corresponding lip movement. The head is then placed on top of the target video at the exact location of the head. In a final step neural rendering is used to render the 3D model in a photorealistic way to produce the final output.

input audio sequence (see image). However, only synchronizing the lip movement results in videos where the head and body movement do not match the content spoken. In the next step we are therefore looking into synthesizing head movements driven by speech. Head poses however, similar to eye blinks or upper body motions, do not correspond one-to-one to the speech (you can say the same sentence with different head movements). Consequently, probabilistic models will be considered for the second part of the project.

**Web-based tool**

While synthetic speech has been largely adopted and accepted by the public (voice assistants, train announcements, GPS navigation), the same is not true for video synthesis technologies. These technologies are still relatively new, and as with any synthesis technology there is the potential for misuse. We therefore believe it is important to have an open exchange with citizens to help shape an understanding for the technologies as well as best practices on how to use these technologies. To kick-off this discussion, the MTC developed an easy to use web tool that allows users to generate synthetic videos using their voice alone. Users can select a moderator and

record a short audio. Our system then generates a synthetic video where the moderator moves his or her lips according to the recorded speech. The tool has been successfully presented at Digitaltag 21 together with SRF at their booth. ■

- > A fully functional pipeline for HD lip sync is ready.
- > Developed an easy to use web tool allowing to generate moderations based on audio recordings.
- > Our web tool was showcased at the Digital Festival 2021 in the booth of SRF to raise awareness about deepfakes.
- > Established collaboration with the Initiative for Media Innovation to develop a prototype for niche sports narrations.

# Supporting article translation

Translating newspaper articles into other languages can immensely increase the number of potential readers. It can, therefore, help to sustainably finance quality journalism. However, article translation is still costly and time-consuming. While there have been significant improvements in automatic translation, manual edits are still required to ensure high-quality texts. In 2021, the MTC built an AI web tool that supports this editing process.

Machine Translation (MT) is the task of automatically translating a text from one language into another. The quality of MT systems has improved dramatically over the past decade. Today, these systems provide quality translations for texts from everyday life and are available through online tools such as Google Translate and DeepL. However, for journalistic texts, the automatic translations still require manual corrections. Because MT systems are typically trained on large datasets of parallel text, they tend to produce general-purpose translations that are not optimized for the writing style of a specific newspaper or domain (economics, politics, sports, etc.).

Over the past months, MTC has been developing an AI tool that helps with this so-called post-editing process of adapting a machine-translated text to better fit a newspaper's requirements. Our AI tool has the potential to reduce the translation time of an article from one language into the other while maintaining the high-quality standards of newspapers.

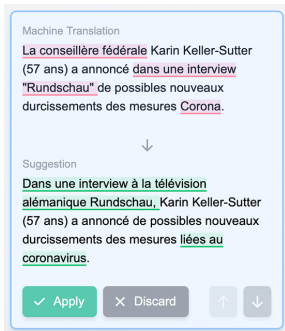
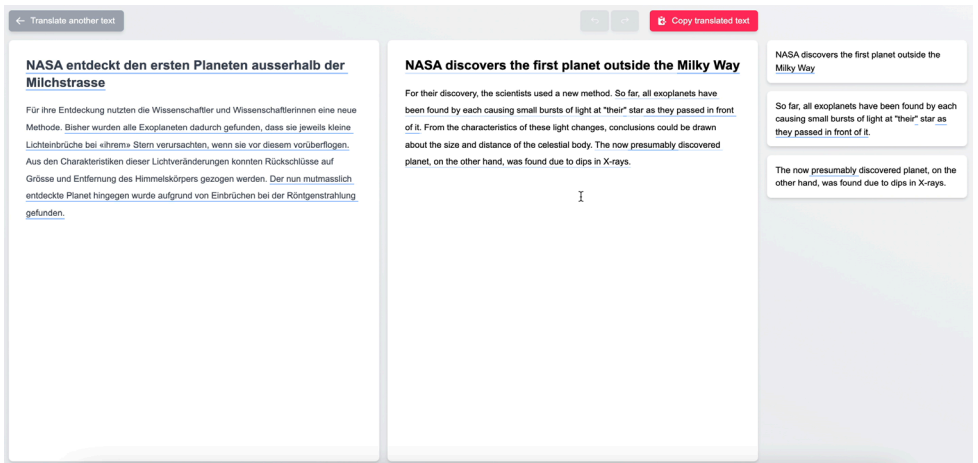
## **Human in the loop**

Our web tool works similarly to other computer-assisted translation tools. After a journal-

ist enters the article to translate, it provides suggestions to edit the machine-translated text (see images). The user can then accept or decline these suggestions as well as directly edit the automatic translation. In addition to this general approach, we want our system to actively work together with journalists and improve the suggestions over time. Therefore, we record user actions and final translation and feed this information back into our AI system. Our approach allows journalists to have direct and complete control over the post-editing process while still benefiting from automatic suggestions. We are planning a study to validate the influence of the tool on the translation speed and quality.

## **Learning from existing newspaper articles**

Suggestions in the web tool are created by a machine learning system. While many newspapers have detailed style guidelines, many conventions about word choices, sentence structures, and writing style are not explicitly stated. Instead of trying to capture these conventions explicitly, our system learns to adapt to the writing style of a newspaper by learning from a large set of news articles. By



[top] Web tool interface.  
 [left] Example: our model learned about "Rundschau" being a German TV show.

and English). This integration allows our system to translate specific Swiss terminology correctly. In addition to TERMDAT, our system supports using any other translation dictionaries available. ■

comparing the machine-translated text to human-translated versions of the exact text, our system can learn the specific changes needed for the translation. We have been working with NZZ, TX Group, Swissinfo, and Keystone-SDA to train our systems for German-English and German-French translations. With the help of the Central Language Services of the Federal Chancellery, we integrated a large part of the TERMDAT translation dictionary into our system. TERMDAT is a multilingual dictionary that contains terminology of Swiss law, public administration, and the public sector in five languages (German, French, Italian, Romansh,

- › Our web-based translation assistant supports journalists with intelligent suggestions while editing translations.
- › Intelligent suggestions are driven by machine learning models trained on large article data sets from NZZ, TX Group, Swissinfo, and Keystone-SDA.
- › Our automatic post-editing models trained on in-domain data for German-French and German-English can already outperform the translation quality of general-purpose MT models.
- › Our post-editing models support the use of translation dictionaries.

# Personalized city-wide augmented reality

In recent years, rapid advances in augmented reality (AR) opened up many new possibilities to incorporate AR into our daily lives. In this project, we built the technology to create a personalized AR experience that seamlessly connects the real world with media content on a city-wide scale.

Augmented reality (AR) allows us to superimpose virtual content on top of the view of the real world. Using sensory and visual information, AR technology can create the illusion of virtual objects and characters that reside in the real world. We believe that there is yet untapped potential for city-wide AR experiences that intelligently connect the real world with augmented content. In this project, we want to demonstrate this potential by building an AR experience that serves the information needs of users exploring a city. Using AR, we can enhance points of interest (POIs) such as restaurants, shopping areas, attraction places with virtual elements and provide users with additional information about these places. However, since there are many potential points of interest in a city, we have to intelligently limit the information shown to the user. To achieve this, we built a context-aware recommender system together with intelligent visualizations to avoid visual overload and provide personalized suggestions of POIs that are predicted to be relevant to users.

## **Augmented points of interest**

Today if you want to find out more information

about a point of interest (POI) such as a restaurant or shop, you typically search for the information on your phone by entering the address or the location's name in a search tool. We want to simplify this process by providing instant access to relevant information. For example, if a user is interested in eating in a restaurant, the restaurant's ratings, type of cuisine, and price range are shown through AR. To achieve this, we built a system that shows relevant information about POIs upon the user's request. Our system highlights places that can provide additional information with rotating icons placed in the real world (see top image on the right). Users can open relevant details on each location with a single touch on the rotating icons.

## **Intelligent city guide**

Today, many services and websites provide user recommendations for items such as clothes, movies, books, restaurants, etc. These recommendations are typically presented as lists for the user to choose from. But how do you provide recommendations in AR? While lists are a natural fit for 2D content on screen, they do not offer a natural way to visualize





The system shows here the restaurant's rating, type of cuisine and reviews.

recommendations in a city-wide setting. To provide recommendations directly in AR, we built a personalized virtual signpost that can give suggestions to the users for places to visit (see image below). Our smart signposts (directions) towards shops and restaurants that fit the users' preferences. A user can bring up the signpost at any time by tapping on the floor. Our recommendation system is intuitively understandable since users know signposts well from everyday life. ■



A virtual signpost gives suggestions to the users.

- > Our smartphone prototype can be used to explore the AR outdoor experience in the Niederdorf area of Zurich.
- > The system can provide instant access to information about points of interest in Zürich.
- > Our AR recommendation models provide contextual recommendations to the users exploring the city.
- > We designed interactive and intelligent visualizations, animations, and different UI types for POIs.
- > According to a live user study and experiments, the initial prototype demonstrates high useability and user satisfaction scores.

# Computer, computer on the wall, what's the fairest of them all?

Images and videos form the visual backbone of journalism. From trailers to teaser images to image galleries, appealing visuals have only grown in importance - especially for online presences. Selecting the best shot from hours of video material or selecting the perfect teaser image from thousands of photos is a complex and time-consuming task. At the MTC, we are teaching computers an understanding of aesthetics to support this selection process.

In fairy tales, magical mirrors can decide about the most beautiful person in the world in seconds. However, automatically analyzing aesthetic features of an image or video remains a challenging problem for computers to solve. Aesthetics and visual appeal are highly subjective. Everyone develops their own sense of style and aesthetics according to personal preferences, experiences, and photography knowledge. Additionally, many different features of an image are considered by experts when selecting key visuals: from the image composition, the mood, and the camera position to colors, depth of field, and facial expressions, many factors are relevant.

While the task is complex, algorithms that detect aesthetic properties can greatly support journalists and video editors in selecting key visuals. Therefore, in this ongoing project, we devised methods for the automatic discovery of aesthetically pleasing images and video shots.

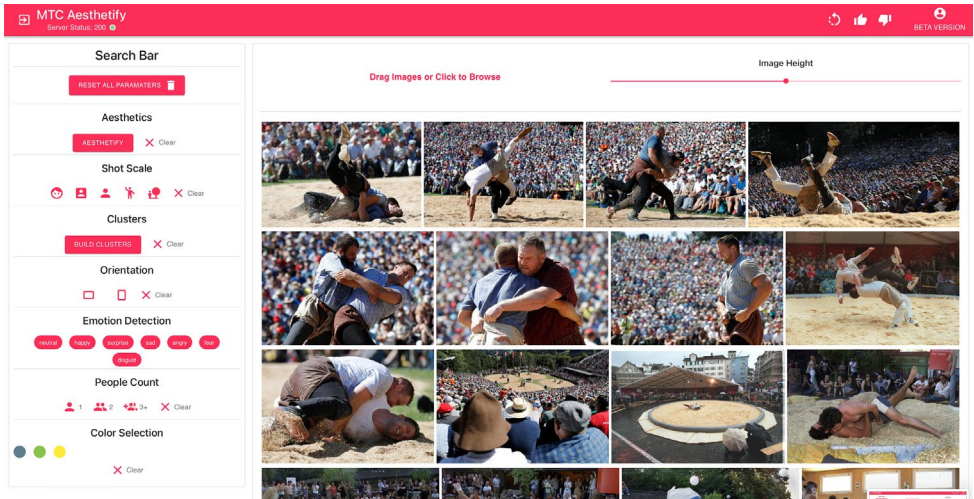
## **Learning from human expertise**

To build an automatic system that can predict

the aesthetic qualities of an image, we have been learning from human knowledge in two different ways. On the one hand, we conducted qualitative interviews with various experts at TX Group and SRF to learn more about the video and image selection process and to understand what features these experts consider relevant for visual aesthetics. On the other hand, we use online photo competitions where photographers upload images for a variety of challenges. Users publicly rate images that enter a challenge in terms of aesthetics. Images with the highest average rating win the challenge. Combining the two aspects, our machine learning system uses the features identified based on the interviews to predict the average rating of images from these online competitions.

## **AI vs. human performance**

We collected an internal test set containing about 500 images ranked manually in terms of aesthetics by five journalists at TX Group. With this test set, we can compare the predictions of our automatic system to human judges, and



The platform's «Aesthetify» feature sorts images according to their predicted aesthetic scores.

we can investigate to what degree humans agree on the aesthetics of an image. Our analysis revealed that our system is currently on par with human performance on this data set. However, the agreement between different journalists, the average journalist and our AI tool is relatively low (<40%), further supporting the idea that aesthetics are subjective even in a clearly defined context.

### Easy-to-use image and video selection tool

We created an image and video shot selection tool that helps users quickly find aesthetically pleasing images and video shots. Users can upload a set of images or a video that is analyzed automatically. Through our web tool, users have access to many features to help them find the perfect shot or image. They can sort images according to predicted aesthetic scores, filter by emotions or cinematic elements, or can enable clustering of similar images. Our prototype system is meant as a showcase for the technology and to get feed-

back from journalists and editors as well as to facilitate user studies on the usefulness of such AI-supported tools. Our industry partners are currently testing the system. ■

- › Collected a gold-standard data set of about 500 manually annotated images by five journalists from TX Group
- › Developed an aesthetic predictor trained on 30K images from image competitions.
- › Developed a pipeline to extract high-level features from images such as shot type, camera angle, and facial emotions
- › An interactive web tool that makes these complex machine learning pipelines easily accessible for testing

# Education

Besides our research projects, we are committed to introducing a young generation of computer science students to the vast possibilities in media technology.

In the past years, significant advances in machine learning and artificial intelligence (AI) have allowed for many novel digital products and have attracted great attention. This recognition can also be seen at ETH, where more students than ever are interested in courses and student projects in the area of machine learning and AI. While these areas have great potential in many fields, we want to introduce students to the unique challenges and opportunities of data-driven products in the news and media industry.

## **Seminar on Media Innovation**

For the third time, we hosted a student seminar for Bachelor students to introduce them to machine learning and AI applications in the field of media technology. Media technology covers many different media types: from text and audio to images and videos and even beyond that to augmented reality experiences. During the fall semester, we gave twenty-three students a glimpse into how data-driven algorithms can be used across all of these formats to create new content and experiences. In the seminar, we are, on the one hand, focusing on recent research papers in computer science. On the other hand, we want to provide insights into how these technologies can solve real-world challenges in the media industry today. David Imseng (recapp), Damian Murezzan (SRF), and Tim Nonner (TX Group) introduced the students to some of the challenges that data scientists and machine learning engineers are solving in media companies in Switzerland today.

## **Student projects**

Our researchers supervised twelve different student projects in 2021. Students programmed new algorithms, experimented with state-of-the-art code, and conducted experiments to validate their ideas in these hands-on projects. Students worked on a variety of topics across our project portfolio. You can find some examples on the right side. ■



### Novel View Synthesis for Personalized Moderations

by Yannick Schmid, Dr. Clara Fernandez Labrador

Computer vision algorithms can create synthetic photorealistic videos of a TV show moderator. These techniques have the potential to individualize video content at scale. Existing techniques however cannot create different

camera views which means the synthetic moderator is always visible from the same camera view. In this project we use a recently introduced technique called Neural Radiance Fields to synthesize new camera views to make synthetic moderations more visually engaging.



### Object Detection and Information Retrieval in Mixed Reality

by Felix Yang, Dr. Saikishore Kalloori

The rapid advances and increasing accessibility of augmented reality (AR) in recent years opened up many new possibilities to incorporate AR into our daily lives. An exciting area for AR is tourism, where one can enhance attractions with virtual elements and provide tourists with additional information about the places they are visiting.

In this project, we developed an AR application that augments various points of interest in Zurich by showing images and facts about them.



### Federated Recommender Systems for News Articles

by Marco Schöb, Dr. Saikishore Kalloori

Recommender systems require a lot of training data to provide good recommendations. This data typically includes user-specific information such as user ratings or click data. To ensure privacy, data ownership and protection

of user's information this data ideally would not need to be shared with a central server to provide recommendations. In this project we use techniques from Federated Learning to train recommender systems without the need for data sharing.

# Outreach

Despite the ever-changing global pandemic situation, the MTC has continued its efforts on outreach by taking part in in-person and virtual events that highlight the results of our research projects.

## Events

[7. Apr] **AI for Media @ SRF INNO Dialog, Zürich** – Presenting an overview of MTC and some project highlights across different media types.

[20. May] **Introduction of MTC @ Digital Society Initiative Community Communication Meeting**  
Introduction of the MTC to the DSI community to explore synergies and potential future collaborations.

[15 June] **Faketaul Reality @ Interaction Design Exhibition, ZHdK, Zürich** – Exhibition created by Paulina Zybinska (ZHdK) and Alessia Paccagnella (ETH) using deepfake technology.

[7. Sep] **AI for Media @ Ringier Feed Your Mind, Zürich** – Presenting an overview of MTC and project highlights across different media types.

[14. Sep] **Pushing boundaries for media technology @ Applied Machine Learning Days, Zürich**  
Keynote during the online business session about how the MTC and its partners are pushing boundaries of what is considered possible in media technology.

[23 Sep] **Moderator Avatar Platform @ Digital Festival, Zürich** – SRF showcased the jointly developed Avatar platform that generates synthetic videos combining the voice of visitors with the faces of famous SRF moderators.

[7. Oct] **Platform for open innovation in media technology @ Meeting of the Leading European Newspaper Alliance (LENA), Zürich** – Talk on how the MTC serves as a platform for open innovation in media technology and how joint research amplifies resources, including some highlights from research projects.

[Oct - Dec] **Moderator Avatar Platform @ ETH Unterwegs** – Together with “ETH Unterwegs” our Avatar platform was part of the projects presented to high school students to peak their interest for stem subjects that can be studied at ETH.

## Scientific Publications

**Stance Detection in German News Articles** *L. Mascarell, T. Ruzsics, C. Schneebeli, P. Schlattner, L. Campanella, S. Klingler, C. Kadar* – Proceedings of the Fourth Workshop on Fact Extraction and VERification 2021 (FEVER 2021)

**Unsupervised Monocular Depth Reconstruction of Non-Rigid Scenes** *A. Takmaz, D. P. Paudel, T. Probst, A. Chhatkuli, M. R. Oswald, L. Van Gool* – International Conference on 3D Vision 2021 (3DV 2021)

**SwissDial: Parallel Multidialectal Corpus of Spoken Swiss German** *P. Dogan-Schönberger, J. Mäder, T. Hofmann* – arXiv preprint arXiv:2103.11401 (2021)

**Horizontal Cross-Silo Federated Recommender Systems** *S. Kalloori, S. Klingler* – Fifteenth ACM Conference on Recommender Systems (RecSys 2021)

**Assessing Evaluation Metrics for Speech-to-Speech Translation** *E. Salesky, J. Mäder, S. Klingler* IEEE Automatic Speech Recognition and Understanding Workshop 2021 (ASRU 2021)

**The 2021 RecSys Challenge Dataset: Fairness is not optional** *L. Belli, A. Tejani, F. Portman, A. Lung-Yut-Fong, B. Chamberlain, Y. Xie, K. Lum, J. Hunt, M. Bronstein, V. W. Anelli, S. Kalloori, B. Ferwerda, W. Shi* – RecSysChallenge 2021: Proceedings of the Recommender Systems Challenge 2021 (RecSys 21)

**Fairness-Aware Engagement Prediction at Scale on Twitter's Home Timeline** *V. W. Anelli, S. Kalloori, B. Ferwerda, L. Belli, A. Tejani, F. Portman, A. Lung-Yut-Fong, B. Chamberlain, Y. Xie, J. Hunt, M. Bronstein, W. Shi* – Fifteenth ACM Conference on Recommender Systems. 2021 (RecSys 21)

**Practical Federated Learning Framework for Small Number of Stakeholders** *C. Schneebeli, S. Kalloori, S. Klingler* – Proceedings of the Fourteenth ACM International Conference on Web Search and Data Mining (WSDM '21)

## MTC in the news

[03 March] **Ein Computer mit «Zürischnurre»** – News article in Beobachter Online.

[07 May] **So soll Siri auf Schweizerdeutsch klingen** – Video report on Blick TV.

[06 Nov] **Können wir bald auf Schwiizerdütsch mit Computern sprechen** – News Article NZZ.

ETH Zürich  
Media Technology Center  
STF  
Stampfenbachstrasse 114  
8092 Zürich

[mtc@ethz.ch](mailto:mtc@ethz.ch)  
[www.mtc.ethz.ch](http://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.