



# Media Technology Center

Annual Report 2023



“By combining the expertise of MTC with the journalistic integrity of Swiss media, we are creating a new paradigm for responsible reporting in the age of AI.”

Dr. Felix Graf

In the beginning was the Word, as stated in the Bible, and in 2023, this concept resonates more than ever. The changes brought forth by Large Language Models (LLMs) are monumental. Accompanying these advancements is a significant increase in the volume of fake news, with hallucinations causing additional uncertainty. Therefore, the role of media in providing context has become critically important.

In the very first issue of the *Neue Zürcher Zeitung* (NZZ) on January 12, 1780, founder Salomon Gessner penned what still holds true today: 'It will not be possible for us to report on worldwide events before they have occurred.' This statement will continue to hold truth in the era of Generative AI. Yet, I firmly believe that generative artificial intelligence (AI) will massively transform our world and the media industry.

Yuval Harari used the term "Oppenheimer Moment" in his 2023 video regarding the developments in Generative AI. Dürrenmatt on the other side beautifully described in his play "The Physicists" that it is impossible to lock away a thought, a technology, or to make things unthought. In the play, Möbius, the physicist who developed the "world formula," attempted to seclude himself in a psychiatric clinic, only to fail miserably.

"If it can be thought, it will be thought." We must find ways to live responsibly with new technologies. As an optimist, I am convinced that technology will be the key to solving some, if not most, of the greatest challenges we face, such as climate change.

In this transformation, the collaboration of Swiss media with ETH Zurich is making a significant contribution. By combining the expertise of one of the world's leading technical universities with the journalistic integrity of Swiss media, we are creating a new paradigm for responsible reporting in the age of AI.

This partnership underscores the necessity of informed and ethical journalism as a cornerstone of democracy and society, especially when navigating the complex landscape of Generative AI. It is not only about reporting the news but also about understanding the implications of these technological advances, deciphering the real from the fabricated, and educating the public on these matters.

As we embrace this new era, we must remember the essence of Gessner's words and the importance of timely, accurate reporting. It is through this lens that we can hope to harness the power of AI for the greater good, ensuring that as we stand on the precipice of this "Oppenheimer Moment," we choose a path that leads to enlightenment and progress rather than to darkness and regression.

Let us, therefore, proceed with caution but also with an unwavering commitment to the truth, so that we may continue to inform and empower society in the face of an ever-evolving technological landscape. The word, after all, remains as potent and as pivotal as it was at the beginning.

Dr. Felix Graf  
CEO at NZZ AG

# People

## Core Team



Alberto Pennino



Armando Schmid



Ayça Takmaz



Daniel Vera Nieto



Fábio Porfírio



Fabio Zünd



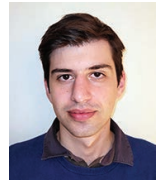
Julien Heitmann



Laura Mascarell



Majed El Helou



Petar Stamenkovic



Piriyakorn  
Piriyatamwong



Ribin Chalumattu



Saikishore Kalloori



Thomas Steinmann



Tianshuai Lu

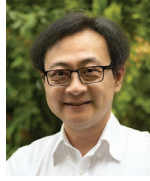
## Industry Experts



Cristina Kadar  
NZZ



Olivier Bühler  
SRF



Zhao Wang  
Ringier AG



Stefan Wabel  
Verband Schweizer  
Medien



Dominic Herzog  
TX Group AG

## Academia Experts



Prof. Andreas  
Krause



Dr. Donald Tillman



Prof. Markus Gross



Nico Lypitkas



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.

# Faithful Summarization of German News

The MTC tackles the persistent challenge of hallucination using LLMs with a focus on German news summarization. The research is crucial in the context of the Swiss media industry to ensure that the models generate accurate information.

Over the past years, there has been remarkable progress in the field of text summarization especially due to the emergent abilities of the recent Large Language Models (LLMs). However, summarization models cannot yet ensure that summaries are factually consistent with the source and may fabricate information (i.e. hallucinate). In addition, most of the research focuses on English, relying on annotated data that is often not available in other languages. More recent works employ multilingual datasets, such as XLSum and mFace, yet they do not include German. Overall, hallucination poses a major issue in journalism, as it is essential to provide the reader with accurate information. Particularly within the Swiss media industry, it is important to be able to detect and mitigate hallucination in German text.

Our project aims to tackle hallucination with a special focus on German news summarization to encourage the research community to implement and evaluate approaches that also consider German. Initially, we focused on hallucination detection and built a manually annotated dataset for this purpose that we

evaluated on multiple open-source LLMs. Our goal for the next phase of the project is to research and experiment with approaches that alleviate hallucination in the output.

## **Absinth: a German Dataset for Hallucination Detection**

To foster research on hallucination in German summarization, the MTC built and open-sourced in collaboration with the University of Zurich the Absinth dataset, a manually annotated dataset for hallucination detection. The dataset consists of ~ 4'300 pairs of article and summary sentences and their corresponding label faithful, intrinsic, or extrinsic hallucination. While intrinsic hallucinations are counterfactual to the source article, extrinsic hallucinations contain information that cannot be verified against the source (see example in Figure 1). The articles come from a publicly available dataset of 20 Minuten news articles, and the corresponding summaries are generated using multiple summarization models, including state-of-the-art pre-trained language models for German summarization and the latest instruction-based LLMs such as OpenAI's GPT-4 and the open-source model

<b>Source:</b> Prof. Park awarded Nobel Prize in Physics.
{F} Nobel Physics Prize goes to Prof. Park.
{I} Prof. Park awarded Nobel Prize in <b>Economics</b> .
{E} Prof. Park <b>(58)</b> awarded Nobel Prize in Physics.

Figure 1: Examples faithful to the source (F), containing an intrinsic hallucination (I), which is counterfactual to the source, and extrinsic hallucination (E), adding new information that cannot be verified.

Faithful
  Intrinsic Hallucination
  Extrinsic Hallucination
  Intrinsic and Extrinsic

#### Summary

Der Rettungsdienst Surselva half Landwirt Hans Müller aus Zürich bei einer Pferdegeburt in Steisslage.

Anschliessend saugte man dem Fohlen das eingeatmete Fruchtwasser aus dem Magen.

Der Rettungsdienst war anlässlich einer Konferenz auf dem Hof anwesend.

Gegen Anfang der Konferenz kam es zum Notfall.

Sowohl Mutterpferd als auch Fohlen sind krank.

### RETTUNGSDIENST SURSELVA HILFT BEI KALBSGEBURT

Eigentlich war der Rettungsdienst Surselva zu einer Weiterbildung bezüglich Mutterkuhherden auf dem Hof von Landwirt Urs Grob in Ilanz GR. Doch als es Komplikationen bei der Geburt eines Kalbs gab, halfen die Sanitäter kurzerhand mit, dieses gesund auf die Welt zu bringen.

Im Kontakt mit Mutterkuhherden kann es zu Verletzungen von beispielsweise Wanderern oder Bikern kommen. Aus diesem Grund absolvierte der Rettungsdienst Surselva am Mittwoch eine Weiterbildung auf der Weide von Landwirt Urs Grob in Ilanz GR.

Figure 2: Screenshot of the annotation interface used to annotate the Absinth dataset. The summary was automatically generated using a summarization model, and the annotation task consists of labelling whether the highlighted summary sentence is faithful to the source article (below) or contains intrinsic, extrinsic hallucinations, or both.

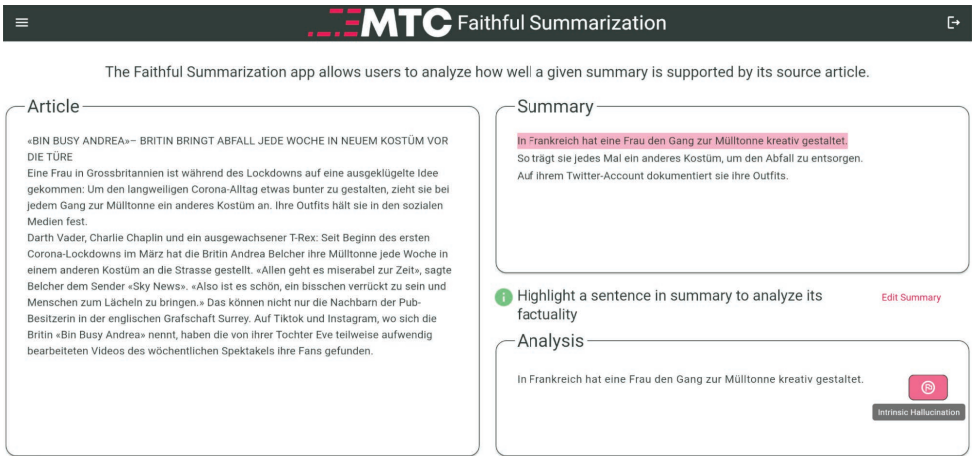


Figure 3: Screenshot of the MTC fact-aware demo application. The interface allows users to analyze the faithfulness of a summary against the given article.

## Llama 2 by Meta.

A team of 12 students from ETH Zurich and the University of Zurich performed the annotation task, achieving a substantial inter-annotator agreement of 0.77 Fleiss' kappa score. To ensure high-quality annotations, participants attended an initial in-person training, and their performance was continuously assessed against a gold standard annotated by domain experts, allowing us to promptly provide them with clarifications on the annotation task if necessary. Additionally, they used an intuitive annotation interface that the MTC adapted to the task (see Figure 2).

## Inconsistency Detection with Absinth and LLMs

LLMs have recently come to the forefront in natural language processing, allowing for significant advancements in the field. These models, which are characterized by an augmented

parameter scale, show a notable improvement in performance compared to well-established pre-trained language models like GPT-2.

Over the last year, the research community built and released multiple open-source LLMs. We therefore evaluate their performance at detecting hallucination using our Absinth dataset on multiple prompting approaches, such as fine-tuning, few-shot, or chain-of-thought prompting. In particular, we chose various model sizes of the Llama 2 family and Mistral, as they show state-of-the-art performance on different benchmarks. Since we focus on the German language, we also consider LeoLM models, which adapt the Llama 2 or Mistral models to German by further pre-training on German data. Our preliminary results show the strong abilities of chain-of-thought prompting, which enhances the model's reasoning by providing a series of intermediate



steps to generate the desired output, and confirm related work findings that the learning abilities of LLMs increase as the model scale becomes larger.

To support journalists in verifying the information from a given article and its corresponding summary, the MTC implemented and deployed a fact-aware demo application that uses our best-performing LLM model fine-tuned on chain-of-thought prompting. More specifically, the application allows users to automatically analyze the faithfulness of individual summary sentences against the source article and obtain more information on the type of hallucination if any (see Figure 3). ■

- › Ensuring accurate information is crucial for the Swiss media industry.
- › We built and open-sourced Absinth, the first dataset for hallucination detection in German.
- › Absinth distinguishes between intrinsic and extrinsic hallucinations at summary sentence level.
- › We evaluated the performance of LLMs at detecting hallucinations.
- › Our experimental results show the strong capabilities of chain-of-thought prompting.
- › We built a demo application that verifies information in summaries using our detection models.
- › The next phase of the project will focus on alleviating hallucination in the generated summaries.

# Follow-up Emotion and Stance Detection for German Text

We built an emotion viewer tool and a stance viewer tool using our labelled dataset and proposed models. We designed an emotion-aware recommender system using the expressed emotions from the article text.

Detecting and extracting information such as emotions and stances from news articles helps news publishers improve user engagement metrics such as next-article click predictions and provides meaningful insights to improve business opportunities. In the previous project, we focused our research on building an annotation tool to enable the annotators to provide a stance label for a given question and articles, and an emotion label for a given article text on paragraph level and full-article level. The outcome was a dataset called 'CHeeSE': The German Dataset of Swiss (CH) News Articles for Stance and Emotion Detection. The dataset has 91 debate questions and 670k Swiss news articles (Blick, NZZ, NZZaS) with 3'693 annotated question-article pairs (Stance, paragraph emotion, article emotion).

In this follow-up project, we aimed to develop end-user prototypes that demonstrate the functionality and possibilities of our emotion and stance detection models. We further explored the usage of emotions for recommender systems where we aimed to incorporate the user and articles expressed emotions into a recommender system model and generate accurate news article recommendations.

## **Design and Development of Prototypes**

For end-user prototypes, we have built two tools. The first one is an emotion viewer tool, which, for a given text from a news article, predicts the emotions that the text conveys. Figure 4 presents a snapshot of the emotion viewer interface. For an input German news article text, the model predicts emotions (we have 8 different emotion types) that are expressed in the text. The second prototype we built is a debate portfolio prototype that displays articles grouped by their stance. Figure 5 presents a snapshot of the stance viewer tool. For example, given the topic "exit nuclear energy", the user can obtain an overview of news articles that are «in favor», «against», or «discussing» the topic.

## **Emotions-Aware News Recommender Systems**

Recommender systems tackle the problem of information overload by predicting news articles that are relevant and interesting to users based on their previous interactions with news articles. In an emotions-aware recommender system, the expressed emotions from news articles are extracted and then these expressed emotions are used in



Figure 4: The UI of Emotion Viewer Tool.

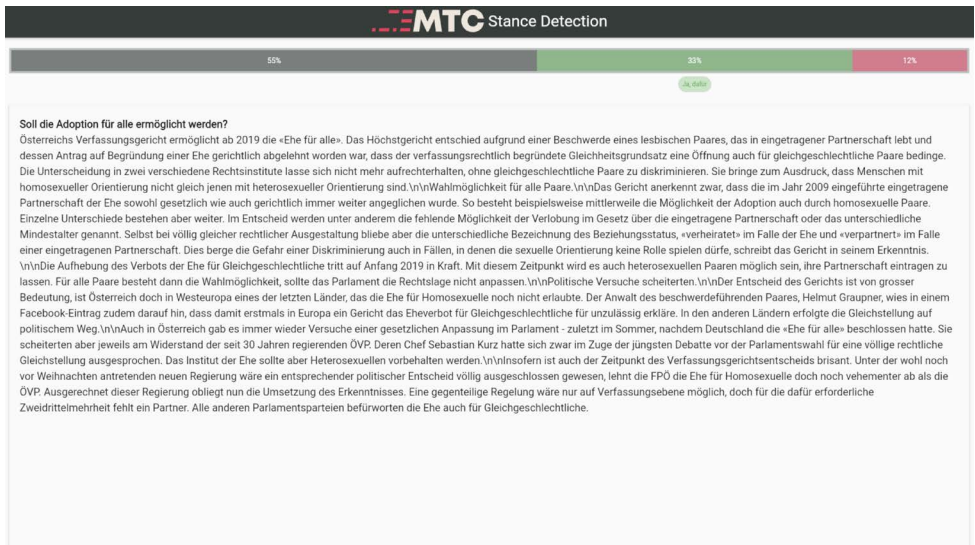


Figure 5: The UI of Emotion Viewer Tool.

the model training and predicting relevant and interesting news articles for users. We trained a predictive model using the above-collected 'CHeeSE' dataset to extract the expressed emotions from the text. We designed three recommendation models: a content-based model, a factorization machines (FM) model, and a deep Neutral FM Model (DeepFM). We conducted experiments using NZZ data with 38K users, 6K news articles, and 7 million user-article impressions.

Our offline experiments showed that the usage of expressed emotions boosts the recommendation performance when compared to ignoring emotions. Our results indicate that emotions contain meaningful information in recommending news articles to users. Our experiments further revealed that the usage of raw (extracted) features is not beneficial, and modelling text and emotions in a linear combination does not yield good performance. This means that if the extracted features are used directly without any further modifications, we notice a drop in the recommendation performance. We need to model emotions intelligently, and higher-order techniques like factorization machines and deep factorization models demonstrated good benefits in ranking performance.

We further conducted experiments to understand when using expressed emotions shows a major benefit or if there are any specific scenarios we could use them. Our results suggest that for users with few ratings available (referred to as a semi-cold start scenario) to the recommender system, exploiting the emotions has shown significant improvement in rank-

ing when compared to a model that does not exploit emotions during recommendation. ■

- › Developed two research prototypes - an emotion viewer tool and a stance viewer tool.
- › Developed a news recommender system to exploit the usage of expressed emotions from the article text.
- › Our experiments demonstrate that expressed emotions boost the recommender system's performance.
- › Our experiments show that emotions can be used to improve recommendations during semi-cold start scenarios.

# Artificial but Natural Voices

We explored the challenge of synthesizing speech from text and developed an AI model that generates expressive German speech from German text. We also adapt the model to produce German speech with a Swiss-German accent, and address the creation of novel unique voices.

The human voice is central to almost every type of presentation. Be it the speaker on television commenting on the pictures of the latest reports, the news presenter on the radio, or a narrator in a podcast. However, automatic text-to-speech systems (TTS) have dramatically improved over the past years and are already a viable alternative to recorded human speech. End-to-end systems relying on deep neural networks and generative adversarial training have produced increasingly human-sounding speech. In this project, our main objective was to develop an open-source method capable of generating expressive unique voices out of text inputs, specifically in German, both with and without a Swiss-German accent.

## Audio Preprocessing

To reach this goal, we have first built a data processing pipeline suitable for audio data samples. Our audio processing pipeline receives raw audio data samples, and performs the following steps: source separation, audio denoising, diarization, transcriptions, and finally embedding and clustering that can be useful for downstream machine learning tasks, analysis, and training. With source separation we decouple the audio into vocal and non-vocal elements. We then perform audio

denoising to remove noise from the retrieved vocal data and enhance the quality of speech. We proceed with diarization: a necessary step of speech separation where each audio is segmented according to the different speakers present.

Subsequently, to train a TTS model, text data is necessary. We perform transcription on the extracted audio samples to obtain the needed text data. We finally extract additional information over the audio segments by performing speaker embedding and clustering to identify single speakers and/or categorical clusters of speakers. Upon surveying the state-of-the-art solutions, whether industrial or academic, we selected the best available open-source TTS method to build on, given the importance of open-sourcing for trust and transparency in media.

## Diffusion-Based Generation Approaches

We explored two different research directions. First, with the help of a student assistant doing an internship with us, we delved into diffusion-based spectrogram generation to integrate into pre-existing vocoder-based TTS pipelines. In TTS literature, a spectrogram is a visual representation of the spectrum of frequencies of an audio signal as it varies with

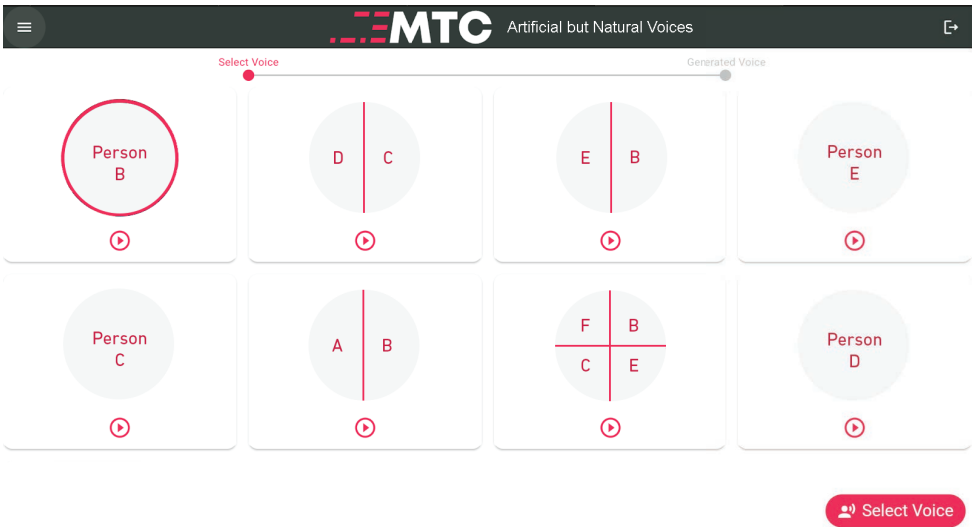


Figure 6: Demo voice selection, between different identities, as well as different mixtures of identities.

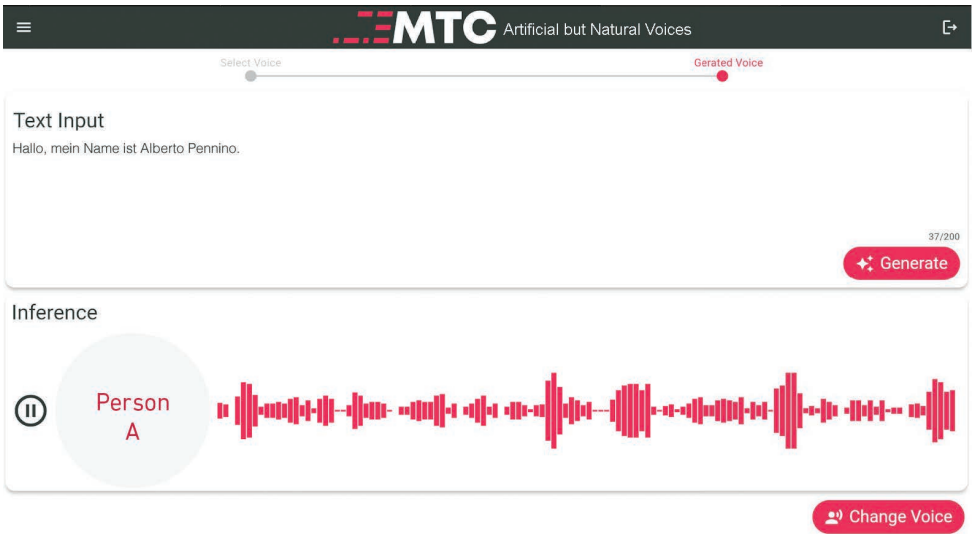


Figure 7: Demo text-to-speech, with a selected identity.

time. A vocoder is a neural architecture used to generate waveform signals from spectrograms. We experimented with the introduction of diffusion-based generative processes within the spectrogram space in the VITS pipeline. Unfortunately, our experiments proved inconclusive due to the strong negative impact of the pre-trained vocoders, which proved to undo any changes introduced by the previous spectrogram generation modifications.

Secondly, we explored full text-to-audio diffusion-based generation. Given the results of our previous experiments, highlighting the limitations of pre-trained vocoders, we decided to substitute the full TTS generative pipeline with a diffusion-based architecture. We tried multiple combinations of diffusion-based generators and diffusion-based vocoders. The models proved to be extremely complex and, unfortunately, never converged to a usable prototype.

### **VITS: Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech**

Ultimately, we obtained pre-trained checkpoints of the VITS model through Meta's open-sourced MMS model, which they made available for the German language. The given checkpoints result from an extensive training procedure on Meta's large compute resources and data: 100'000 training steps using eight V100 GPUs on a dataset of 44'700 hours. We adapted the checkpoint model through further training on Swiss-German accent data: a smaller dataset of approx. 10 hours of moderations recorded by our industry partners.

We then proposed mixed-identity data training

to obtain novel unique voices for our text-to-speech models. Until the end of the project, we will analyze the various facets of our model and results with a user study.

We finalized our project demo prototype for users to easily test and try out our final models. In our final prototype, users can select existing pre-trained identities and listen to voice snippets as shown in Figure 6. Once a voice has been selected, as presented in a snapshot in Figure 7, the user can proceed to write any text and the pre-trained VITS model will generate the corresponding voice. ■

- › We researched state-of-the-art TTS methods as well as emerging startups in the field.
- › We developed a comprehensive audio processing pipeline for managing, cleaning, and enhancing audio datasets.
- › We discovered a pre-trained, open-source German TTS model featuring expressive voices.
- › We utilized Swiss data to fine-tune our TTS model for the German language, incorporating a Swiss-German accent.
- › We introduced a latent interpolation strategy through mixed-data training to generate unique voices.
- › We are currently conducting a user study to analyze various aspects of the project results.

# Education

In addition to research, the MTC is dedicated to teaching the next generation of computer science students about the exciting possibilities of media technology.

## Seminar on Media Innovation

During the autumn semester 2023, for the fifth time, the MTC hosted the Seminar on Media Innovation tailored specifically for Bachelor students, which saw the enthusiastic participation of 24 motivated individuals. This seminar is designed with two goals in mind: first, to explore the latest advances in media technology, and second, to hone the student's presentation and analytical skills essential for navigating the dynamic landscape of innovation and research.

Each participant was tasked with presenting a seminal paper, fostering an environment of scholarly discourse and knowledge exchange among peers. Students explored a range of topics from recommender systems, fact-checking, AI-driven animation, music and image synthesis, to AI models such as GANs, Transformers, BART, and BERT.

## Bachelor's and Master's Theses

During 2023, the MTC staff oversaw six student projects, fostering practical experiences for students to create new algorithms, test cutting-edge code, validate their concepts, and carry out user studies. The project topics encompassed recommender systems, summarization, emotion analysis, online advertising, and argument extraction.

## ZHdK Interdisciplinary Module

Students from different departments at the Zürcher Hochschule der Künste (ZHdK) signed up to attend an interdisciplinary 3-week module that requires them in a team of 2 to 4 students to design and develop a product, an experience, or an idea related to research and technologies developed at the MTC. The three resulting projects were "Woven Wonders - AR adventures on an enchanted carpet" (Figure 8), "chatter.box - solve dilemmas together" (Figure 9), and "leARn - a language learn app" (Figure 10).

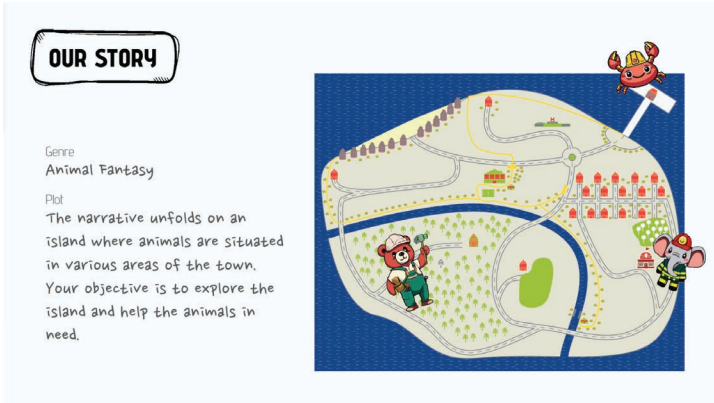
## Courses of Prof. Ryan Cotterell

In 2023, Prof. Ryan Cotterell taught the following courses at the Department of Computer Science: Advanced Formal Language Theory, Large Language Models, Understanding Context-Free Parsing Algorithms, Philosophy of Language and Computation II, Natural Language Processing, and Neural Networks and Computational Complexity. Over 400 students have participated in these lectures. ■

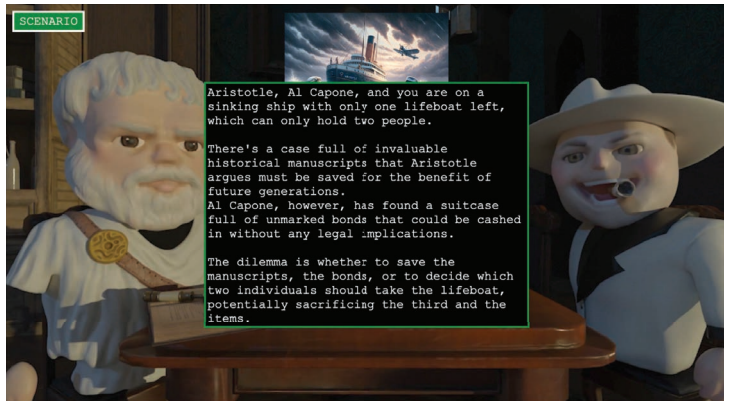


All student projects can be accessed on our website.





[left] Figure 8: "Woven Wonders - AR adventures on an enchanted carpet".



[right] Figure 9: "chatter.box - solve dilemmas together".



[left] Figure 10: "leARn - a language learn app".

# Outreach

Engaging in conferences and events to showcase the outcomes of our research projects is a key aspect of the MTC's ongoing outreach efforts.

## Events

- 16-19 January **ENTER 2023 Conference @ Johannesburg, South Africa** – Paper presentations of “Towards Recommender Systems in Augmented Reality for Tourism,” and of “Talking Houses: Transforming Touristic Buildings into Intelligent Characters in Augmented Reality”.
- 11 May **SwissMediaForum @ Lucerne, Switzerland** – Presentation of “Development of new technologies for the Swiss Media”.
- 6 June **“Künstlich & Intelligent?” @ SRF-Fernsehstudio, Zurich, Switzerland** – Live demo booth of the “Audio-Driven Video Synthesis of Personalized Moderations” project, and presentation of the MTC.
- 23-24 June **The Joint Computation + Journalism European Data & Computational Journalism Conference 2023 @ ETH Zurich, Switzerland** – Conference organization and hosting.
- 9-14 July **ACL 2023 @ Toronto, Canada** – Poster Presentation of “OpenTIPE: An Open-source Translation Framework for Interactive Post-Editing Research”.
- 23-27 July **ACM SIGIR 2023 @ Taipei, Taiwan** – Paper presentation of “A retrieval system for images and videos based on aesthetic assessment of visuals”.
- 2-7 September **Intelligent Content Creation Tools for AR/VR Workshop @ Seoul, South Korea** – Workshop held as part of a research collaboration between Switzerland and South Korea funded by Innosuisse and KIAT.

- 11-15  
September **INLG 2023 @ Prague, Czech Republic** – Paper Presentation of “Entropy-based Sampling for Abstractive Multi-document Summarization in Low-resource Settings”.
- 18-22  
September **ACM RecSys @ Singapore** – Co-Organizer of RecSys 2023 Challenge, paper presentations of “RecSys Challenge 2023 Dataset: Ads Recommendations in Online Advertising”, and of “RecSys Challenge 2023: Deep Funnel Optimization with a Focus on User Privacy”.
- 2-6 October **ICCV 2023 @ Paris, France** – Paper presentation of “3D Segmentation of Humans in Point Clouds with Synthetic Data”, poster presentation, and workshop organization.
- 8-11 October **IEEE ICIP 2023 @ Kuala Lumpur, Malaysia** – Paper Presentation of “Fuzzy-Conditioned Diffusion and Diffusion Projection Attention Applied to Facial Image Correction”.
- 24 October  
10 November **ZHdK Module @ ZHdK and MTC, Zurich** – Students at the Zürcher Hochschule der Künste created prototypes and demos based on research at the MTC.
- 14-17  
November **ACM SIGGRAPH MIG 2023 @ Rennes, France** – Poster Presentation of “Artwalks via latent diffusion models”.
- 11-16  
December **NeurIPS 2023 @ New Orleans, LA, USA** – Paper presentation of “OpenMask3D: Open-Vocabulary 3D Instance Segmentation”.

# Outreach

## Scientific Publications

### **Towards Recommender Systems in Augmented Reality for Tourism**

*Saikishore Kalloori, Ribin Chalumattu, Felix Yang, Severin Klingler and Markus Gross*  
30th Annual International eTourism Conference (ENTER23), Johannesburg, South Africa

### **Talking Houses: Transforming Touristic Buildings into Intelligent Characters in Augmented Reality**

*Saikishore Kalloori, Ribin Chalumattu, François Chalet, Martin Zimper, Severin Klingler and Markus Gross*  
30th Annual International eTourism Conference (ENTER23), Johannesburg, South Africa

### **OpenTIPE: An Open-source Translation Framework for Interactive Post-Editing Research**

*Fabian Landwehr, Thomas Steinmann and Laura Mascarell*  
61st Annual Meeting of the Association for Computational Linguistics (ACL 2023), Toronto, Canada

### **A Retrieval System for Images and Videos based on Aesthetic Assessment of Visuals**

*Daniel Vera Nieto, Saikishore Kalloori, Fabio Zünd, Clara Fernandez Labrador, Marc Willhaus, Severin Klingler and Markus Gross*  
SIGIR '23: The 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, Taipei, Taiwan

### **Entropy-based Sampling for Abstractive Multi-document Summarization in Low-resource Settings**

*Laura Mascarell, Ribin Chalumattu and Julien Heitmann*  
16th International Natural Language Generation Conference (INGL 2023), Prague, Czech Republic

### **RecSys Challenge 2023 Dataset: Ads Recommendations in Online Advertising**

*Rahul Agrawal, Sarang Brahme, Sourav Maitra, Abhishek Srivastava, Athirai Irissappane, Yong Liu and Saikishore Kalloori*  
ACM Recommender Systems Challenge Workshop 2023 (RecSysChallenge), Singapore

**RecSys Challenge 2023: Deep Funnel Optimization with a Focus on User Privacy**

*Rahul Agrawal, Sarang Brahme, Sourav Maitra, Saikishore Kalloori, Abhishek Srivastava, Yong Liu and Athirai Irissappane*

17th ACM Conference on Recommender Systems (RecSys 2023), Singapore

**3D Segmentation of Humans in Point Clouds with Synthetic Data**

*Ayça Takmaz, Jonas Schult, Irem Kaftan, Mertcan Akçay, Bastian Leibe, Robert W. Sumner, Francis Engelmann and Siyu Tang*

19th IEEE/CVF International Conference on Computer Vision (ICCV 2023), Paris, France

**Fuzzy-Conditioned Diffusion and Diffusion Projection Attention Applied to Facial Image Correction**

*Majed El Helou*

30th IEEE International Conference on Image Processing (ICIP 2023), Kuala Lumpur, Malaysia

**OpenMask3D: Open-Vocabulary 3D Instance Segmentation**

*Ayça Takmaz, Elisabetta Fedele, Robert W. Summer, Marc Pollefeys, Federico Tombari and Francis Engelmann*

37th Annual Conference on Neural Information Processing Systems (NeurIPS 2023), New Orleans, LA, USA



All of MTC's scientific publications can be accessed on our website.

ETH Zürich  
Media Technology Center  
STF  
Stampfenbachstrasse 114  
8092 Zürich

Contact:  
+41 44 832 82 86  
fabio@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.