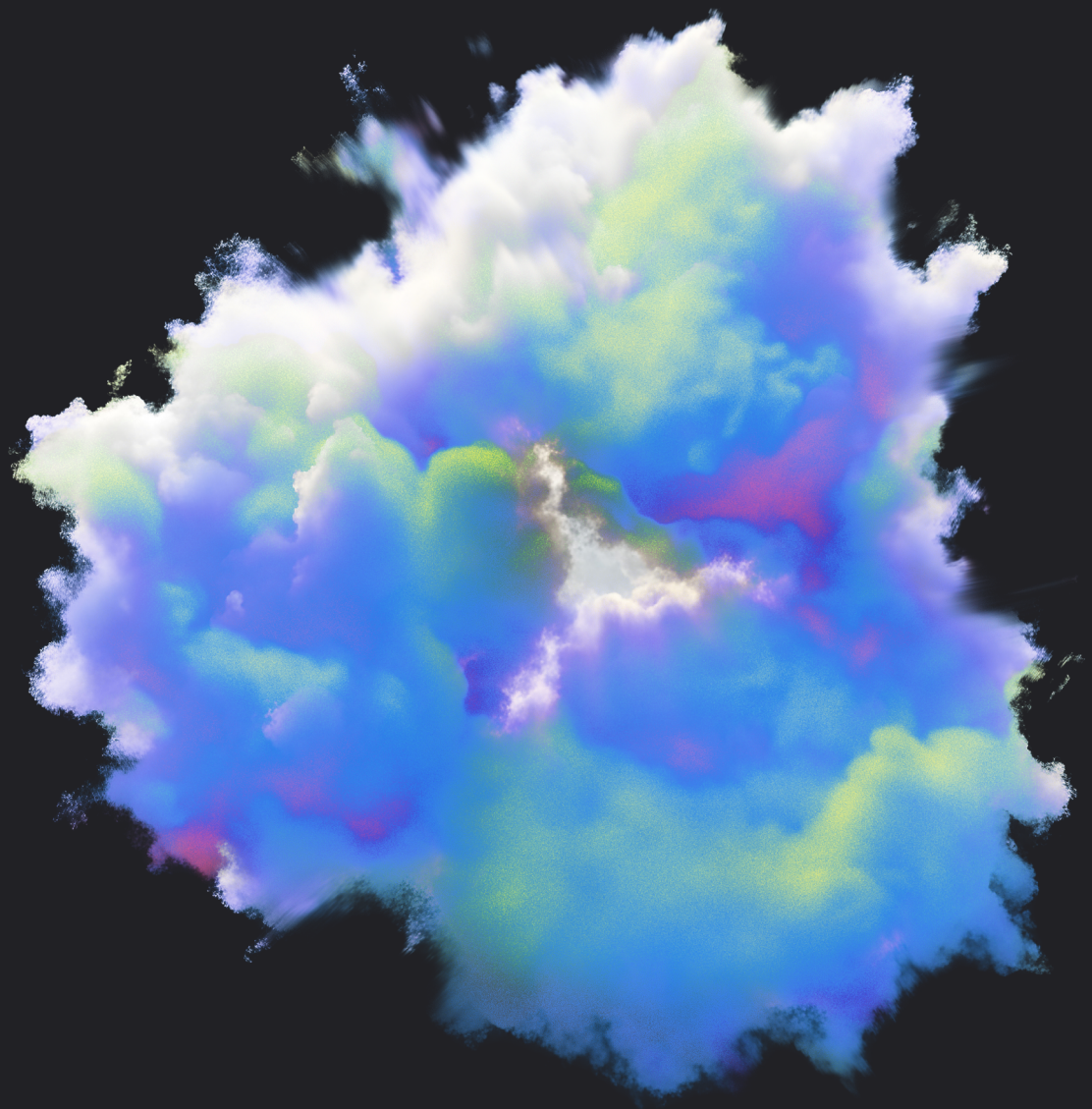
SOUNDS OF THE METAVERSE

AI IN SOUND & MUSIC



FOREWORD

SOUNDS OF THE METAVERSE

THE AIM OF THE CONFERENCE IS TO BRING TOGETHER AN INTERNATIONAL COMMUNITY OF RESEARCHERS AND PRACTITIONERS WORKING WITH SOUND IN VIRTUAL ENVIRONMENTS, INCLUDING (AMONG OTHERS) GAME ENGINES. THE CONFERENCE FOLLOWS A ONE-DAY FORMAT CONSISTING OF TALKS AND PRESENTATIONS OF WORKS BY PEOPLE FROM DIVERSE BACKGROUNDS WHO ENGAGE WITH VARIOUS ASPECTS OF WORKING WITH SOUND IN DIGITAL AND VIRTUAL ENVIRONMENTS, AS WELL AS BY SCHOLARS EXAMINING VIRTUAL ENVIRONMENTS FROM A CULTURAL STUDIES PERSPECTIVE.

THIS YEAR'S EDITION WILL PARTICULARLY ADDRESS TOPICS RELATED TO THE USE OF ARTIFICIAL INTELLIGENCE IN THE FIELDS OF SOUND AND MUSIC.

DURING THE CONFERENCE, ATTENDEES WILL HAVE THE OPPORTUNITY TO LISTEN TO IMMERSIVE SOUND WORKS RECOGNIZED IN THE EIGHTH EDITION OF THE STUDENT 3D AUDIO PRODUCTION COMPETITION, ORGANIZED ANNUALLY BY THE INSTITUTE OF ELECTRONIC MUSIC AND ACOUSTICS, A UNIT OF THE UNIVERSITY OF THE ARTS IN GRAZ.

THE OPENING LECTURE OF THE CONFERENCE, TITLED "THREE PERSPECTIVES ON MUSIC AI," WILL BE DELIVERED BY DR. ENG. MATEUSZ MODRZEJEWSKI, AN ASSISTANT PROFESSOR AT WARSAW UNIVERSITY OF TECHNOLOGY, A RESEARCHER INTO THE APPLICATIONS OF AI IN MUSIC, AND A PERFORMING MUSICIAN.

WE SINCERELY HOPE YOU WILL FIND THE CONFERENCE PROGRAM BOTH INSPIRING AND ENJOYABLE. AS WE NAVIGATE THE HYPE SURROUNDING THIS YEAR'S MOST PROMINENT BUZZWORD, OUR GOAL IS THAT THESE SESSIONS WILL PROVIDE YOU WITH A GROUNDED, PROFESSIONAL PERSPECTIVE - MOVING BEYOND THE TRENDS TO UNCOVER THE TRUE CREATIVE AND TECHNICAL POTENTIAL OF THESE TOOLS.

SCIENTIFIC COMMITTEE:

DR PRZEMYSŁAW DANOWSKI

DR MARTA GOSPODAREK

DR MATEUSZ MODRZEJEWSKI

ORGANIZING COMMITTEE:

DR PRZEMYSŁAW DANOWSKI

DR MARTA GOSPODAREK

VOLUNTEERS:

DANIEL ADAMCZYK

ALEKSANDRA BIELECKA

IGA KAMIŃSKA

JULIA PIEGZA

FILIP ROSZKOWSKI

JUSTYNA ŚMIETANA

CONFERENCE PROGRAM

10:00–10:20

OPENING REMARKS

DR PRZEMEK DANOWSKI
CHOPIN UNIVERSITY OF MUSIC

10:20–11:00

KEYNOTE SPEECH: THREE PERSPECTIVES ON MUSIC AI

DR INŻ. MATEUSZ MODRZEJEWSKI
WARSAW UNIVERSITY OF TECHNOLOGY

11:00–11:30

PRESENTATION & LIVE ACT: GENERATIVE ALGORITHMS AS AN EXTENSION OF PRIMITIVE SOUNDS

MGR GNIEWOMIR TOMCZYK

11:30–12:00

ARTISTS + ALGORITHMS: CO-CREATING RESPONSIBLE AI IN MUSIC

MGR, INŻ. WOJCIECH HAZANOWICZ
MUZAIC.AI

12:00–12:30

UPSCALING HRTFS WITH GENERATIVE AI

DR INŻ. PAULA PIETRZAK
WARSAW UNIVERSITY OF TECHNOLOGY

12:30–13:30

LUNCH BREAK

26.11.2025

KAROL SZYMANOWSKI AUDITORIUM
CHOPIN UNIVERSITY OF MUSIC
WARSAW

SOUNDS OF THE METAVERSE

CONFERENCE PROGRAM

13:30–14:30
LISTENING SESSION
PRESENTATION OF THE AWARD-WINNING ENTRIES FROM THE 9TH EDITION OF THE STUDENT 3D AUDIO PRODUCTION COMPETITION (S3DAPC).

14:30–15:00
WHAT IS ZAIKS LAB AND HOW CAN IT HELP YOU?
AGNIESZKA CICHOCKA (ZAIKS)

15:00–15:30
FUTURE FRAGMENTS
MGR MARCIN NOWICKI
(NOVIKI STUDIO)

15:30–16:00
COFFEE BREAK

16:00–16:30
DETECTING DEEPFAKE SPEECH WITH CONVOLUTIONAL NEURAL NETWORKS
INŻ. MATEUSZ ZYCH
(WARSAW UNIVERSITY OF TECHNOLOGY)

16:30–17:00
THE USE OF AI IN WORKING WITH DIALOGUES IN MOVIES
DR IWONA EJSMUND, MGR ANNA BALCEWICZ, MGR MICHAŁ JAKUBIAK
(CHOPIN UNIVERSITY OF MUSIC, THE FELIKS NOWOWIEJSKI ACADEMY OF MUSIC IN BYDGOSZCZ)

17:00–17:30
CAN DEEP NEURAL NETWORKS RECOGNIZE EMOTIONS FROM POLISH SPEECH?
INŻ. KATARZYNA STEC
(WARSAW UNIVERSITY OF TECHNOLOGY)

17:30–18:30
WORKSHOP: GENERATIVE AUDIO MODELLING WITH TRANSFORMERS
DANIEL FARONBI
(MUSIC AND AUDIO RESEARCH LABORATORY, NEW YORK UNIVERSITY)

18:30–18:45
CONFERENCE CLOSING

SOUNDS OF THE METAVERSE

THREE PERSPECTIVES ON MUSIC AI

MATEUSZ MODRZEJEWSKI, PHD

This talk explores generative and analytical aspects of audio and music AI, explained in plain English. Together, we will examine the intersection of art and technology through the lens of neural networks and consider the impact of these technologies on music, presented from three distinct perspectives.

MATEUSZ MODRZEJEWSKI, PHD - SOFTWARE ENGINEER, RESEARCHER, CONFERENCE SPEAKER, AUTHOR AND CO-AUTHOR OF PAPERS ON MUSIC INFORMATION RETRIEVAL AND AUDIO AI. ASSISTANT PROFESSOR AT THE INSTITUTE OF COMPUTER SCIENCE OF WARSAW UNIVERSITY OF TECHNOLOGY, WHERE HE LEADS AN AUDIO INTELLIGENCE LAB. PREVIOUSLY AT APPLE (MUSIC MACHINE LEARNING TEAM, APPLE MUSIC). HAS ALSO WORKED WITH RESEARCH AND ENGINEERING TEAMS OF OTHER FORTUNE 500 COMPANIES, PROVIDING AI SOLUTIONS AND ANALYTICS.

APART FROM HIS SCIENTIFIC AND ENGINEERING WORK, HE IS ALSO AN EXPERIENCED TOURING MUSICIAN, HAVING PERFORMED FOR AUDIENCES OF UP TO 150,000 PEOPLE AND HAVING TOURED IN POLAND, CHINA, VIETNAM, THE UK, GERMANY, UKRAINE, LITHUANIA AND ESTONIA, AMONG OTHERS. SOME OF THE ARTISTS HE HAS PLAYED WITH INCLUDE THE DUMPLINGS, GRUBSON, MAREK DYJAK, CHŁOPCY KONTRA BASIA, MARIA SADOWSKA, PABLOPAVO I LUDZIKI, MAJKA JEŻOWSKA, MICHAŁ MILCZAREK TRIO.

GENERATIVE ALGORITHMS AS AN EXTENSION OF PRIMITIVE SOUNDS

GNIEWOMIR TOMCZYK

Concert: solo performance during the lecture (improvisation and an excerpt from the radioplay "LemAI", along with animation) on a drum set with generative AI algorithms.

GNIEWOMIR TOMCZYK: DRUMMER, PRODUCER, AND CREATOR OF THE FUTUROPOLIS PROJECT EXPLORES HOW ARTIFICIAL INTELLIGENCE AND GENERATIVE ALGORITHMS ENABLE NEW FORMS OF EXPRESSION IN DRUMMING — EXPANDING THE NATURAL SOUND OF THE INSTRUMENT AND REDEFINING THE RELATIONSHIP BETWEEN THE MUSICIAN, SOUND, AND THE ROLE OF THE INSTRUMENT WITHIN THE CONCERT SPACE.

ARTISTS + ALGORITHMS: CO-CREATING RESPONSIBLE AI IN MUSIC

WOJTEK HAZANOWICZ
MUZAIC

What does responsible, artist-centered music AI look like in practice? In this talk I'll share how a Poland-based start-up built models that composed 9.5 million soundtracks in a year—while keeping artists, legality, and reliability at the core. We'll follow "Julia," a digital specialist who needs fast, affordable, risk-free music for video: why bespoke tracks are slow, stock is frustrating, and most AI still raises legal and quality concerns. I'll show how Muzaic delivers editable structures (sections, loops, variations) in a single file, stable audio quality, and ethically sourced training (no third-party songs). We'll explore integrations with editors and generative video tools, and why hyper-personalized sound at scale is the next frontier—done with consent, credit, and compensation in mind.

WOJTEK HAZANOWICZ, FOUNDER AND CEO AT MUZAIC, IS A PROFESSIONAL MUSICIAN, ENGINEER, AND ENTREPRENEUR. HIS ADVENTURE WITH MUSIC BEGAN WITH CLASSICAL MUSIC AND LATER MOVED INTO THE AREAS OF INNOVATION AND ARTISTIC EDUCATION. AS PRESIDENT OF JM POLAND AND SECRETARY OF THE POLISH MUSIC COUNCIL, HAZANOWICZ HAS BEEN INVOLVED IN THE DEVELOPMENT OF MUSIC CULTURE FOR YEARS. HE, IN TURN, FOUNDED HIS FIRST BUSINESS (IN THE AREA OF IT AND E-COMMERCE) DURING HIS COMPUTER SCIENCE STUDIES (COMPUTER SCIENCE, WARSAW UNIVERSITY OF TECHNOLOGY) IN 2006.

UPSCALING HRTFS WITH GENERATIVE AI

AGNIESZKA PAULA PIETRZAK, PHD

Personalized head-related transfer functions (HRTFs) are an important component of creating high-quality immersive audio scenes, but often require dense, time-consuming measurements. This work explores a data-efficient alternative - upscaling sparse, low-resolution HRTFs to full-sphere resolution with generative models, and benchmarking them against classical interpolation. Using public HRTF databases, the high-resolution HRTFs are reconstructed from a small subset of directions and assessed with objective metrics and perceptual localization models. The study shows that generative upscaling can recover spatial cues more faithfully than interpolation, especially in left–right consistency, with some remaining challenges in front–back errors.

AGNIESZKA PAULA PIETRZAK, PHD - RESEARCH AND TEACHING ASSISTANT PROFESSOR AT THE DEPARTMENT OF ELECTROACOUSTICS IN THE INSTITUTE OF RADIOELECTRONICS AND MULTIMEDIA TECHNOLOGY, WARSAW UNIVERSITY OF TECHNOLOGY. SHE CONDUCTS SCIENTIFIC WORK AND HAS EXPERIENCE IN IMPLEMENTING RESEARCH PROJECTS IN THE FIELD OF DIGITAL AUDIO SIGNAL PROCESSING, SPATIAL SOUND, CONDUCTING HEARING TESTS, AND SPEECH PROCESSING. AUTHOR OF SCIENTIFIC AND POPULAR SCIENCE PUBLICATIONS, ACTIVELY PARTICIPATING IN POLISH AND INTERNATIONAL CONFERENCES AND TRAINING. MEMBER OF THE BOARD OF THE POLISH SECTION OF THE AUDIO ENGINEERING SOCIETY, MEMBER OF THE EUROPEAN ACOUSTICS ASSOCIATION, AND THE POLISH ACOUSTICAL SOCIETY.

CATEGORY 1:
CONTEMPORARY MUSIC / COMPUTER MUSIC

Gold: NATALIA QUINTANILLA CABRERA
(composer, sound engineer)
Affiliation: University of Washington, Seattle (USA)
Piece: **"Abejas"**
Duration: 8'36"

Silver: JAKOB GILLE
(composer, sound engineer)
Affiliation: University of Music and Performing Arts Graz (Austria)
Piece: **"Self-Motion Study"**
Duration: 10'52"

Bronze: JIAJING ZHAO
(composer, sound engineer)
Affiliation: University of the Arts London (UK)
Piece: **"Fomo"**
Duration: 7'11"

CATEGORY 2:
AUDIO DRAMA / DOCUMENTARY / SOUNDSCAPES

Silver: THEODOR-ALEXANDRU PETREA
(composer, sound engineer)
Affiliation: SRH Berlin University of Applied Sciences (Germany)
Piece: **"Desmoterion"**
Duration: 3'49"

Bronze: DAVID ADLBERGER
(composer, sound engineer)
Affiliation: University of Music and Performing Arts Graz (KUG) / FH Joanneum (Austria)
Piece: **"Self"**
Duration: 3'59"

CATEGORY 3:
MUSIC RECORDING / STUDIO PRODUCTION

Gold: ELISA POLOV
(composer, sound engineer)
Affiliation: St. Pölten University of Applied Sciences (Austria)
Piece: **"Don't Go Near the Edge"**
Duration: 3'12"

Bronze: MARCIN STUDNIARZ
(sound engineer)
Affiliation: Chopin University of Music (Poland)
Piece: **"Kuldahar"** (comp. Jeremy Soule)
Duration: 2'41"

LISTENING SESSION

THE STUDENT 3D AUDIO PRODUCTION COMPETITION (S3DAPC) IS A PREMIER INTERNATIONAL PLATFORM DEDICATED TO IMMERSIVE AUDIO. OPEN TO STUDENTS WORLDWIDE, IT CHALLENGES PARTICIPANTS TO CREATE SPATIAL SOUND WORKS USING AMBISONICS TECHNOLOGY. ENTRIES SPAN THREE CATEGORIES: CONTEMPORARY/COMPUTER MUSIC, AUDIO DRAMA/DOCUMENTARY, AND STUDIO PRODUCTION. THE COMPETITION SERVES AS A VITAL BRIDGE BETWEEN ACADEMIA AND THE PROFESSIONAL WORLD, WITH FINALISTS PRESENTING THEIR WORK AT THE PRESTIGIOUS TONMEISTERTAGUNG (TMT).

S3DAPC IS A COLLABORATIVE INITIATIVE LED BY THE INSTITUTE FOR ELECTRONIC MUSIC AND ACOUSTICS (IEM) AT THE UNIVERSITY OF MUSIC AND PERFORMING ARTS GRAZ (AUSTRIA), CONCEIVED BY MATTHIAS FRANK AND FRANZ ZOTTER. IT IS ORGANIZED IN PARTNERSHIP WITH THE VERBAND DEUTSCHER TONMEISTER E.V. (VDT), WHICH HOSTS THE AWARDS, AND SUPPORTED BY ORF MUSIKPROTOKOLL AND SOUNDING FUTURE.

WHAT IS ZAIKS LAB AND HOW CAN IT HELP YOU?

AGNIESZKA CICHOCKA

ZAiKS.Lab – a space for experiment, education, and cooperation between creators and technology. It is a place where innovation, copyright, and creativity meet in the spirit of responsible cultural development.

AGNIESZKA CICHOCKA IS A CULTURAL INNOVATION STRATEGIST AND CREATIVE TECH EXPERT WITH EXTENSIVE EXPERIENCE BUILDING INTERNATIONAL PROGRAMS AT THE INTERSECTION OF ART, TECHNOLOGY, AND ENTREPRENEURSHIP. SHE HAS DELIVERED ARTISTIC AND INNOVATION PROJECTS ACROSS EUROPE AND NORTHERN AFRICA AND SERVES AS AN EXPERT AT THE EUROPEAN INSTITUTE OF TECHNOLOGY CULTURE & CREATIVITY.

CURRENTLY, SHE PLAYS A CENTRAL ROLE AS HEAD OF ACCELERATION AT ZAIKS LAB, WHERE SHE BUILDS INNOVATION PATHWAYS FOR THE CREATIVE AND CULTURAL SECTOR, SUPPORTS FOUNDERS, AND HELPS SHAPE THE FUTURE OF CREATIVE TECH SOLUTIONS WITHIN THE MUSIC AND CULTURAL RIGHTS ECOSYSTEM. SHE IS ALSO THE CO-AUTHOR OF THE 2025 CREATIVE-TECH INDUSTRY REPORT FOR POLAND, CREATED IN PARTNERSHIP WITH FUELARTS AND THE CENTER FOR CREATIVE INDUSTRIES DEVELOPMENT.

FUTURE FRAGMENTS
MARCIN NOWICKI

Future Fragments is a platform and network of reflection dedicated to exploring contemporary and future phenomena in digital culture, technology, and artificial intelligence. The initiative investigates the concept of intelligence as a relational and emergent phenomenon—arising through interactions between organisms, systems, and environments.

Rather than viewing technology as the opposite of nature, Future Fragments perceives it as another layer of the planet's ecosystem—coexisting with biological and cultural processes. The project creates an open, decentralized space for the exchange of ideas, where voices of researchers, artists, designers, and thinkers from diverse fields come together to examine what intelligence—both human and nonhuman—is becoming in the post-digital era.

Through lectures, discussions, workshops, presentations, and publications, Future Fragments expands the discourse on the future of technology and its influence on humanity, culture, and the planet. It is an initiative for researchers, creators, and anyone seeking a deeper understanding of the dynamic transformations occurring at the intersection of technology, art, and consciousness.

MARCIN NOWICKI IS THE CO-FOUNDER OF NOVIKI – A CONCEPTUAL POST-STUDIO THAT TRANSCENDS THE BOUNDARIES OF TRADITIONAL DESIGN DISCIPLINES BY MERGING DESIGN, ART, AND RESEARCH ON TECHNOLOGY. IN HIS PRACTICE, HE EXPLORES THE IMPACT OF EMERGING TECHNOLOGIES ON CREATIVITY, VISUAL PERCEPTION, AND CONTEMPORARY CULTURE.

AS THE BMW ART CLUB ARTIST 2024, HE CO-CREATED DREAM OF THE MACHINE, A PROJECT COMBINING STAGE ART, INSTALLATION, AND VIDEO ESSAY WITH ARTIFICIAL INTELLIGENCE. HIS INTERDISCIPLINARY APPROACH TO DESIGN AND REFLECTION ON THE ROLE OF TECHNOLOGY IN CULTURE LED TO THE CREATION OF FUTURE FRAGMENTS – A RESEARCH AND ARTISTIC PLATFORM AVAILABLE AT FUTUREFRAGMENTS.ONLINE.

MARCIN NOWICKI HAS SERVED AS AN EXTERNAL ADVISOR AND LECTURER AT THE ROYAL ACADEMY OF ART (KABK) IN THE NETHERLANDS, AND AS A GUEST SPEAKER AT UNIVERSITIES IN TAIWAN, INCLUDING NATIONAL TAIWAN NORMAL UNIVERSITY AND NATIONAL TAIPEI UNIVERSITY OF TECHNOLOGY, AS WELL AS ACROSS EUROPE. HE IS CURRENTLY A PHD CANDIDATE AT SWPS UNIVERSITY, WHERE HE TEACHES COURSES ON DESIGN IN PUBLIC SPACE. HIS STUDIO HAS RECEIVED NUMEROUS DISTINCTIONS, INCLUDING STGU STUDIO OF THE YEAR AND A GOLD MEDAL AT THE EUROPEAN DESIGN AWARDS. HE HAS COLLABORATED WITH INTERNATIONAL INSTITUTIONS AND RESEARCH GROUPS SUCH AS CHELSEA COLLEGE OF ARTS IN LONDON AND DESIGN INQUIRY, PRESENTING HIS WORKS IN ART GALLERIES, PUBLIC SPACES, AND VIRTUAL ENVIRONMENTS.

DETECTING DEEPFAKE SPEECH WITH CONVOLUTIONAL NEURAL NETWORKS

MATEUSZ ZYCH, MGR INŻ

In recent years, the detection of synthetic speech has become one of the key challenges in the field of cybersecurity and speech processing systems. Significant advances in voice cloning techniques now allow the generation of realistic-sounding recordings that imitate a specific person's speech, which facilitates spoofing attacks and identity theft. This presentation will focus on the phenomenon of cloned speech, the mechanisms behind generating deepfake audio, and modern detection methods based on machine learning. Special attention will be given to convolutional neural network–based approaches, as well as a practical example using an original model for cloned speech classification.

MGR INŻ. MATEUSZ ZYCH - A GRADUATE IN TELECOMMUNICATIONS WITH A SPECIALIZATION IN TELEINFORMATICS AND CYBERSECURITY FROM THE WARSAW UNIVERSITY OF TECHNOLOGY. PRESIDENT OF THE STUDENT ELECTROACOUSTICS ASSOCIATION AT THE WARSAW UNIVERSITY OF TECHNOLOGY, ACTIVELY ENGAGED IN ACADEMIC ACTIVITIES. EXPERIENCED IN PROJECTS RELATED TO SPATIAL AUDIO, SPEECH BIOMETRICS, AUTOMATIC SPEAKER VERIFICATION SYSTEMS, AS WELL AS CYBERSECURITY. PARTICIPANT OF INTERNATIONAL SCIENTIFIC CONFERENCES, PRESENTING RESEARCH IN THE FIELDS OF ACOUSTICS AND SPEAKER VERIFICATION SYSTEMS. MEMBER OF THE AUDIO ENGINEERING SOCIETY.

THE USE OF AI IN WORKING WITH DIALOGUES IN MOVIES

This presentation examines current applications of AI in movie dialogue editing. It begins with an overview of modern noise-reduction and audio-restoration plugins, including AI-driven tools capable not only of suppressing noise but also of deriving impulse responses directly from recorded material. The second part focuses on ElevenLabs and its use in generating, shaping, and extending dialogue with synthetic voices. The session concludes with an interactive test in which participants attempt to identify whether the presented voice samples are AI-generated or spoken by a real actor, illustrating both the strengths and limitations of contemporary AI audio technologies.

IWONA EJSMUND, PHD

- SOUND ENGINEER AND ACADEMIC LECTURER HOLDING A PHD DEGREE. SHE IS ASSOCIATED WITH THE CHOPIN UNIVERSITY OF MUSIC IN WARSAW, WHERE SHE PROVIDES ARTISTIC SUPERVISION FOR STUDENT PROJECTS AT THE DEPARTMENT OF SOUND ENGINEERING.

IN HER PROFESSIONAL WORK, SHE COMBINES TECHNICAL EXPERTISE WITH MUSICAL SENSITIVITY. SHE SPECIALIZES IN SOUND POST-PRODUCTION FOR FILMS, PARTICULARLY TELEVISION AND DOCUMENTARY PROJECTS, AND HAS EXTENSIVE EXPERIENCE WITH A WIDE RANGE OF SERIES, INCLUDING THE CIPHER FORTRESS MYSTERY, WAR GIRLS, THEATRE OF TELEVISION PRODUCTIONS, AND DOCUMENTARIES. MOST RECENTLY, THE FILM THE GREY AMATEUR DIRECTED BY PETRO ALEKSOWSKI, ON WHICH SHE WORKED, RECEIVED SEVERAL AWARDS AT INTERNATIONAL FILM FESTIVALS. SHE IS CURRENTLY WORKING ON THE SERIES THE BLACK DEATH DIRECTED BY KUBA CZEKAJ, WHERE SOUND PLAYS A KEY, EXPERIMENTAL ROLE — ALWAYS ENSURING HER WORK REMAINS DEEPLY ALIGNED WITH THE FILM'S ARTISTIC VISION. SHE SPECIALIZES IN SOUND QUALITY AND AESTHETICS IN MEDIA.

THE USE OF AI IN WORKING WITH DIALOGUES IN MOVIES

ANNA BALCEWICZ

COMPLETED THE SOUND ENGINEERING IN MULTIMEDIA PROGRAM UNDER JOANNA NAPIERALSKA, PHD HABIL., ALONG WITH AN ADDITIONAL SPECIALIZATION IN RADIO SOUND ENGINEERING IN THE CLASS OF ANDRZEJ BRZOSKA, PHD HABIL. SHE CURRENTLY WORKS AT POLISH RADIO, WHERE SHE DIRECTS AUDIO DRAMAS FOR PROGRAM ONE. HER RECENT WORK INCLUDES THE SOUND DIRECTION FOR ESCAPE FROM GRANDE MÉNESTRANDISE BY ANNA URSZULA KUCHARSKA, FEATURED ON AN ALBUM COMBINING MUSICAL COMPOSITIONS WITH THIS ORIGINAL RADIO PLAY.
ANNA'S ARTISTIC INTERESTS FOCUS ON THE RELATIONSHIP BETWEEN SOUND, MUSIC AND NARRATIVE, PARTICULARLY WITHIN CONTEMPORARY FORMS OF AUDIO DRAMA.

MICHAŁ JAKUBIAK

IS A SOUND ENGINEER AND LECTURER AT THE FELIKS NOWOWIEJSKI ACADEMY OF MUSIC IN BYDGOSZCZ. HE COMPLETED HIS MASTER'S DEGREE AT THE FRYDERYK CHOPIN UNIVERSITY OF MUSIC, SPECIALIZING IN SOUND ENGINEERING FOR FILM AND TELEVISION. HE IS A TWO-TIME LAUREATE OF THE AES STUDENT RECORDING COMPETITION IN THE SOUND FOR VISUAL MEDIA CATEGORY, WINNING THE GOLD AWARD IN MADRID AND THE SILVER AWARD IN NEW YORK.

CAN DEEP NEURAL NETWORKS RECOGNIZE EMOTIONS FROM POLISH SPEECH?

KATARZYNA STEC, INŻ.

While Automatic Speech Emotion Recognition (ASER) has been a subject of intense research for many years, comprehensive solutions specifically designed for the Polish language remain scarce. This paper presents the evaluation results of four deep neural network models that demonstrate potential in recognizing emotions spontaneously expressed in Polish. The study examines two multilingual ASER models and two models originally developed for the Russian language, selected due to its linguistic proximity to Polish within the Slavic group. The analysis was conducted on the BES corpus, a database comprising natural (spontaneous) Polish emotional utterances. The results indicate that the highest-performing model achieved an accuracy of approximately 59%. These findings underscore the inherent challenges in detecting subtle and ambiguous emotional expressions in real-world conditions and highlight the necessity for further research dedicated to developing models optimized specifically for the Polish language.

INŻ. KATARZYNA STEC - A GRADUATE WITH A BACHELOR'S DEGREE IN TELECOMMUNICATIONS, SPECIALIZING IN WIRELESS AND MULTIMEDIA TECHNOLOGIES, FROM THE WARSAW UNIVERSITY OF TECHNOLOGY. SHE COMPLETED HER THESIS AT THE ELECTROACOUSTICS DIVISION OF THE INSTITUTE OF RADIOELECTRONICS AND MULTIMEDIA TECHNOLOGY, CONDUCTING RESEARCH ON THE EFFECTIVENESS OF DEEP NEURAL NETWORK MODELS FOR SPEECH EMOTION RECOGNITION IN POLISH. HER SCIENTIFIC INTERESTS FOCUS ON AUDIO SIGNAL PROCESSING. SHE IS CURRENTLY PURSUING A MASTER'S DEGREE IN TELECOMMUNICATIONS, SPECIALIZING IN SOUND AND IMAGE ENGINEERING, FURTHER EXPANDING HER KNOWLEDGE AND RESEARCH SKILLS. SHE ACTIVELY PARTICIPATES IN SCIENTIFIC AND POPULAR-SCIENCE CONFERENCES DEDICATED TO ACOUSTICS AND SOUND ENGINEERING.

WORKSHOP: GENERATIVE AUDIO MODELLING WITH TRANSFORMERS

DANIEL FARONBI

The transformer is a neural network architecture that has risen to prominence because of its efficiency in modelling language. Recent advances in audio tokenization approaches has allowed the transformer to be used to model audio in a discrete fashion similar to language. In this talk, we will go over the design of a transformer as well as the audio tokenization process. We will discuss introductory concepts in machine learning and build up to advanced audio models. We will look at the history of token based audio generation and discuss various models that implement this approach. Finally, we will show step by step how to train a generative transformer on a very simple dataset for those interested in learning how to work with generative models.

DANIEL FARONBI IS A MUSIC TECHNOLOGY PHD CANDIDATE IN THE MUSIC AND AUDIO RESEARCH LABORATORY (MARL) AT NEW YORK UNIVERSITY. HIS RESEARCH INTERESTS LIE IN CONTROLS FOR GENERATIVE AUDIO SYSTEMS. HE HAS SERVED AS A RESEARCH INTERN FOR TOP AUDIO COMPANIES LIKE DOLBY LABORATORIES AND BOSE CORPORATION AND HOLDS TWO BACHELOR'S DEGREES IN MUSIC TECHNOLOGY AND COMPUTER ENGINEERING WITH A MINOR IN MATHEMATICS FROM THE UNIVERSITY OF NEBRASKA. DANIEL IS ALSO AN ACTIVE JAZZ PIANIST, MUSIC PRODUCER, AND COMPOSER. HE HAS PERFORMED FOR MUSIC GROUPS SPANNING A WIDE VARIETY OF GENRES AND HAS RECEIVED BOTH JAZZ AND CLASSICAL TRAINING.

HE HAS WORKED AS A STUDIO MANAGER FOR STUDIO 381 AT THE UNIVERSITY OF NEBRASKA AND HAS PUBLISHED TWO ALBUMS. DANIEL HAS PERFORMED AT WORLD RENOWNED JAZZ VENUES LIKE THE DJANGO AND NUBLU AND MAINTAINS AN ACTIVE PERFORMANCE SCHEDULE IN NYC.

UMFC

Uniwersytet Muzyczny
Fryderyka Chopina

Chopin University of Music
Okólnik 2 St.
00–368 Warsaw
info@chopin.edu.pl