RESEARCH INSTITUTE

SCIENTIFIC REPORT 2024

Contents

Introduction

Founded in 1991, the Idiap Research Institute is an independent, non-profit research foundation based in Martigny, Switzerland. Idiap is a Research Institute of National Importance that was established by the State of Valais, the City of Martigny, the École Polytechnique Fédérale de Lausanne (EPFL), the University of Geneva, and Swisscom. Idiap is dedicated to advancing research, innovation, and education in Artificial Intelligence (AI).

Idiap aims to leverage the potential of AI to support Switzerland's societal and industrial needs, seize opportunities for innovation, and address global challenges. To achieve these objectives, the Institute has established four research programs (RPs) designed to foster collaboration and knowledge exchange among its research groups, while facilitating cross-disciplinary research. Research on AI Fundamentals supports all RPs, namely Human-AI Teaming (RP1), which develops innovative AI solutions that empower people to work effectively with AI systems, enabling effective human-AI collaboration and co-creation; Sustainable and Resilient Societies (RP2), which develops solutions to anticipate and mitigate future disruptions, promote sustainability, and combat disinformation; AI for Life (RP3), which leverages AI to advance biological understanding, develop personalized medicine, and improve health outcomes; and AI for Everyone (RP4), which fosters inclusive AI development, empowering individuals from all backgrounds to contribute to and benefit from AI-driven solutions.

AI Fundamentals

- Why are neural networks so effective?
- How can LLMs learn interpretable abstractions?
- What properties of natural language do LLMs really know?
- Progressive graph neural networks with differential privacy guarantees.
- Ethical awareness and controlled inference in language models.
- Bayesian fine-tuning of pre-trained models.

RP1: Human-AI Teaming

- Semantics and social gaze attribute inference in multi-person gaze understanding.
- Can gaze understanding improve the accuracy of head gesture recognition?
- Advancing the understanding of hand manipulations in the wild.
- Combining end-user programming & shared autonomy for human-robot collaboration.
- Frugal robot learning with tensor networks.
- Implicit shape representations for optimal control.
- Geometric algebra for robotics.

RP2: Sustainable and Resilient Societies

- District heating network simulation.
- ChatGPT for biometrics.
- How synthetic datasets can expose real identities.
- Design of imaging targets to characterize thermal cameras and imaging conditions.

RP3: AI for Life

- Artificial networks can learn biological processes.
- Compressive sensing for cardiac microscopy.
- Suppressing noise disparity in training data.
- Spontaneous speech in automatic pathological speech detection.
- Pathological speech detection adaptation in noisy environments.
- Adversarial attacks for automatic pathological speech detection.
- Parkinson's disease speech analysis and detection.
- Creating clinically interpretable AI models.
- Fine-tuning foundation models without sufficient data increases demographic biases.
- Evidence-based natural language inference in oncology.
- What can we learn about autism severity from long, untrimmed videos?

RP4: AI for Everyone

- Cultural diversity in large language models.
- Improving model generalization in mobile sensing.
- Improving fairness without demographic metadata.
- Bias mitigation in face recognition.
- Generative AI Literacy: twelve defining competencies.
- Social context from smartphone data across countries and daily life moments.
- Interpretable image privacy classification.

The above achievements are complemented by the details of externally funded projects, publicly available code and data distributed to the broader research community, and the list of published research articles.

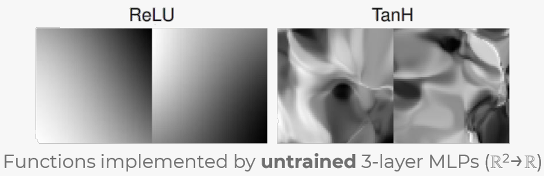
In 2024, Idiap ran 92 externally funded projects. Furthermore, the Institute is active in technology transfer, collaborates with various companies, and creates and supports start-ups through its incubator IdeArk S.A. Idiap pursues an active policy towards reproducible science, with the release of 59 open-source software libraries in 2024. Idiap publishes data through FAIR (Findable, Accessible, Interoperable, Reusable) repositories. The Idiap website offered 27 new publicly available datasets in 2024, for a total of 108.

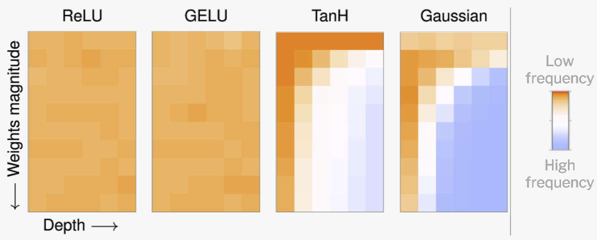
The compilation in this report offers an overview of Idiap's research activities undertaken during the year and showcases the outcomes of collective research efforts.

AI Fundamentals

Why are neural networks so effective?

This work explores fundamental questions in machine learning and offers a fresh perspective on why deep learning has been so successful across a variety of tasks. Indeed, neural networks almost always generalize better to novel data than other types of machine learning models. While most explanations of this success focus on the learning and optimization process, we argue that many properties of neural networks can be explained more fundamentally. In particular, some critical choices in their, such as ReLU activations, makes them likely to fit their training data with simple, low-frequency functions, a phenomenon we termed "*Neural Redshift*". This property is akin to the inductive principle of Occam's razor. Remarkably, this property can be observed even in untrained, random-weight models. This insight is relevant to Idiap's *Human-AI Teaming* research program as it will ultimately help characterize the reliability of AI systems.

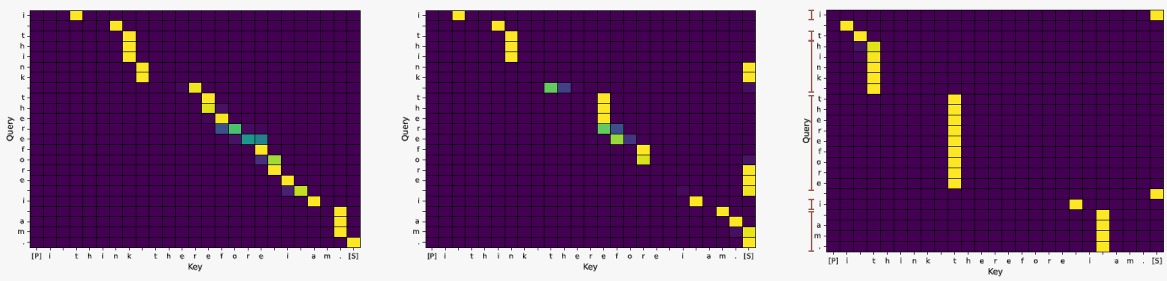




D. Teney, A. M. Nicolicioiu, V. Hartmann, and E. Abbasnejad, Neural redshift: Random networks are not random functions, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

How can LLMs learn interpretable abstractions?

LLMs use the transformer architecture and its layers of attention over many vectors to model the complex nature of language. Different vectors seem to represent different parts of the input, and different layers seem to capture different levels of abstraction, but these relationships are unclear and imprecise. We have developed a new version of attention which pushes transformers towards discovering meaningful abstractions. The layers close to the input (left figure) use many vectors to represent small segments of input, and deeper layers (right figures) learn to use fewer vectors which each represent larger, more abstract units. Our initial results indicate that these abstractions facilitate the interpretability of transformers' learned representations, such as inducing words from character sequences (rightmost figure). We have recently shown that this new version of attention can also be used to accurately reinterpret off-the-shelf pretrained LLMs, showing that we can initialise our models with LLMs pretrained on large amounts of data. The next step is fine-tuning these LLMs with our methods to clarify what abstractions they have learned. These advances will facilitate *Human-AI Teaming* by improving our ability to understand, control and learn from LLMs.

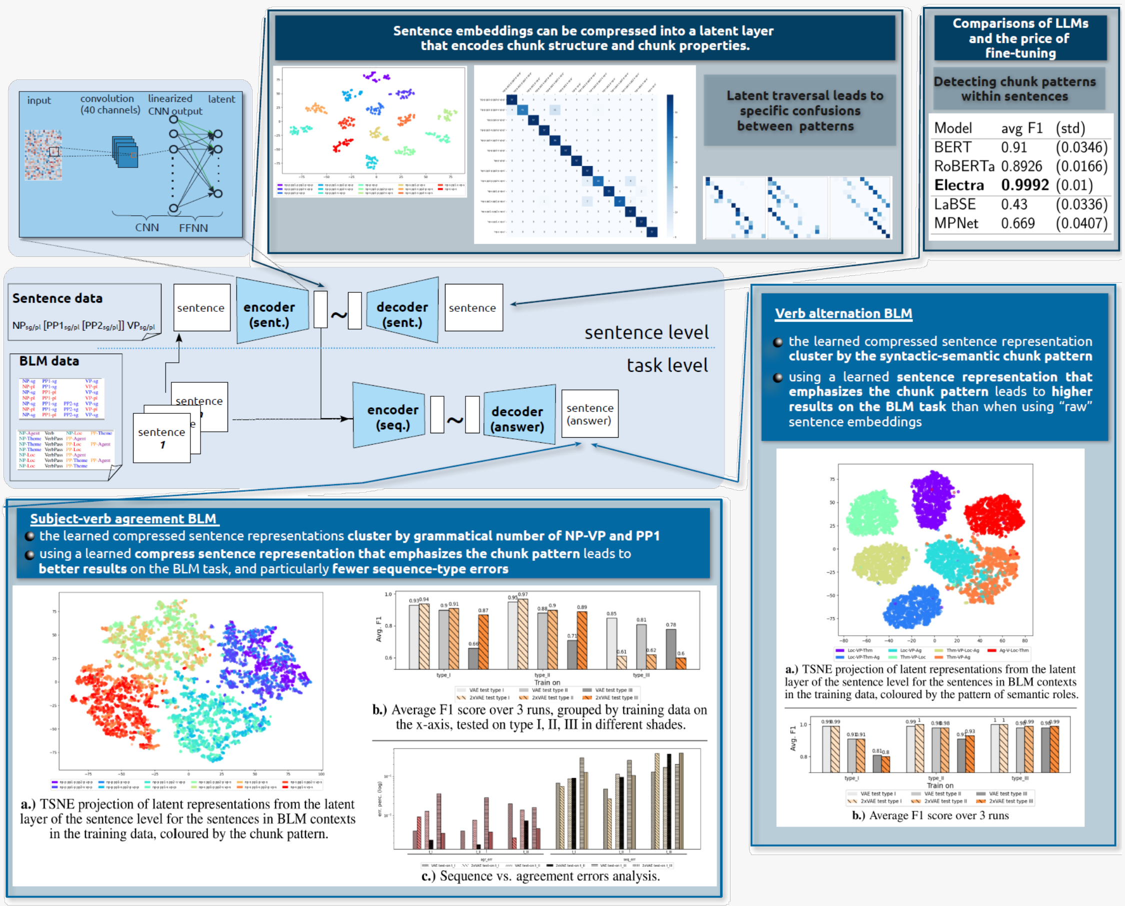


M. Behjati, F. Fehr and J. Henderson, Learning to Abstract with Nonparametric Variational Information Bottleneck, *EMNLP Findings*, 2023.

F. Fehr and J. Henderson, Nonparametric Variational Regularisation of Pretrained Transformers, *Conference on Language Modeling (COLM)*, 2024.

What properties of natural language do LLMs really know?

To develop a system with near-human language capabilities, we need to understand current systems' generalisation and compositional abilities. To address this, we have created a new task in multiple languages that requires both linguistic and reasoning skills. Each task requires identifying relevant linguistic entities and the generative rules that connect them. These richly-structured datasets are then used to probe sentence embeddings generated by transformer models. We investigate whether structural information – specifically, information about chunks and their structural and semantic properties – can be detected in these representations. Through an approach involving indirect supervision, analyses of task performance, and analyses of learned representations, we show that information about chunks and their properties can be obtained from sentence embeddings. Moreover, by using these datasets consisting of sentences with known structure, we test to what degree information about their chunks, such as grammatical number, or semantic role, can be localized in sentence embeddings. Our results show that such information is not distributed over the entire sentence embedding, but is rather encoded in specific regions Understanding how the information from an input text is compressed into sentence embeddings helps understand current transformer models and help build future explainable neural models.



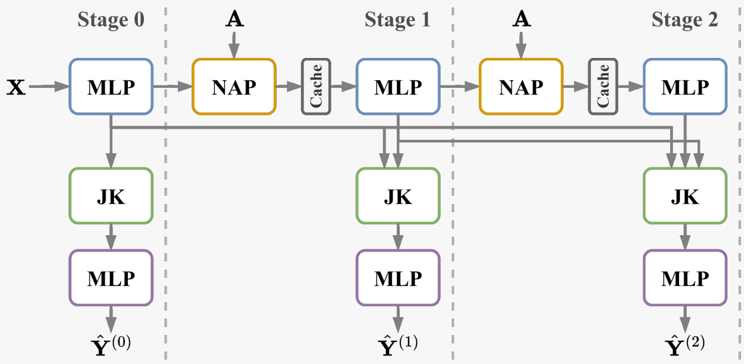
V. Nastase and P. Merlo, Are there identifiable structural parts in the sentence embedding whole? *Blackbox NLP Workshop: Analyzing and Interpreting Neural Networks for NLP*, 2024.

V. Nastase and P. Merlo, Tracking linguistic information in transformer-based sentence embeddings through targeted sparsification, *Workshop on Representation Learning for NLP*, 2024.

Progressive graph neural networks with differential privacy guarantees

Graph Neural Networks (GNNs) have become a popular tool for learning on graphs. GNNs learn from the structural connectivity of graphs by iteratively updating node embeddings through information aggregation and transformation from neighboring nodes. This makes them well suited to node classification, graph classification, and link prediction. However the practical use of GNNs —when applied to human-generated data— can raise privacy concerns, as graph data might contain personal or sensitive information. Differentially-private GNNs have been proposed to preserve privacy while allowing learning over graph-structured datasets. However, reaching a balance between accuracy and privacy in GNNs remains challenging due to the intrinsic structural connectivity of graphs.

We proposed ProGAP, a differentially-private GNN that uses progressive training to improve the accuracy/privacy trade-off. Combined with an aggregation perturbation technique to ensure differential privacy, ProGAP splits a GNN into a sequence of overlapping sub-models that are trained progressively, expanding from the first sub-model to the complete one. ProGAP provides edge-level and node-level privacy guarantees both during training and inference.

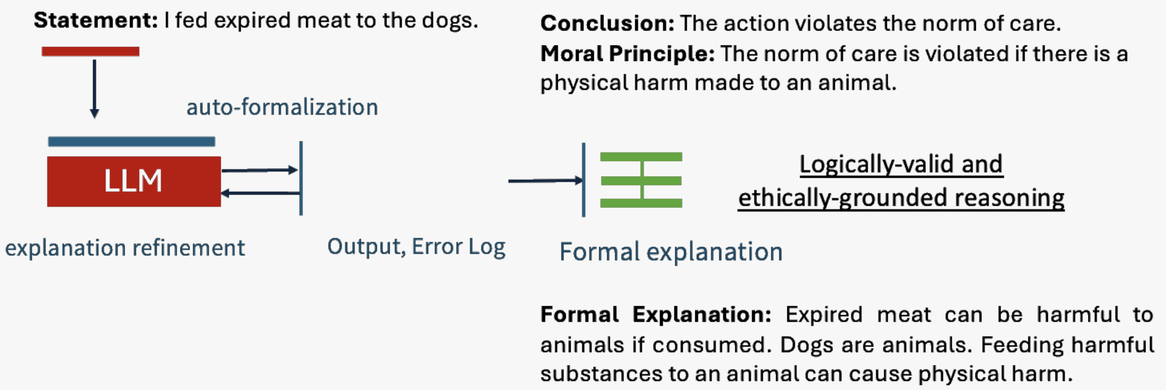


ProGAP architecture's three stages. MLP: Multi-Layer Perceptron; JK: Jumping Knowledge module; NAP: normalized-aggregate-perturb module. Training is done progressively across the three stages, each using its own head MLP. The final prediction is obtained from the head MLP of the last stage.

S. Sajadmanesh and D. Gatica-Perez, ProGAP: Progressive Graph Neural Networks with Differential Privacy Guarantees, *ACM International Conference on Web Search and Data Mining (WSDM)*, 2024.

Ethical awareness and controlled inference in language models

The widespread adoption of Large Language Models (LLMs) is delivering major opportunities for facilitating automated inference, and are progressively making their way into critical application areas such as health, law and policy-making. AI models based on LLMs require mechanisms for supporting ethically-grounded inferences. In this line of work, we are building the foundations to deliver AI models which are capable of more controlled and ethically-aware natural language inference. At the center of the proposed approach is the combination of the complementary properties of content-rich material inference provided by LLMs with logically-valid inferences provided by symbolic solvers. Empirical analyses demonstrate that the close dialogue between these two types of systems can deliver AI systems which are capable of more controlled, logically-valid and grounded explanations.



X. Quan, M. Valentino, L. A. Dennis, and A. Freitas, Enhancing ethical explanations of large language models through iterative symbolic refinement, *Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

X. Quan, M. Valentino, L. A. Dennis, and A. Freitas, Verification and refinement of natural language explanations through llm-symbolic theorem proving, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024.

Bayesian fine-tuning of pre-trained models

What was once referred to as “adaptation” in the field of speech processing is now known as “fine-tuning” in the context of large foundation models. This process has become completely generic. We asked the question “*Is there a right way to do fine-tuning?*” given that generic additional training can be prone to issues such as catastrophic forgetting. Our answer to the question is that fine-tuning can be viewed as a Bayesian process, where the model acts as the prior and the fine-tuning data serves as additional evidence. A further contribution of this work is to demonstrate that both the LoRA and Kronecker methods for fine-tuning can be viewed as Bayesian techniques. This clarifies the relation between these recent methods and classical regularisation techniques such as Elastic Weight Consolidation (EWC). We also developed an effective evaluation strategy that we applied on NLP tasks (language modeling) as well as speech synthesis tasks.

H. Chen and P. N. Garner, Bayesian parameter-efficient fine-tuning for overcoming catastrophic forgetting, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2024.

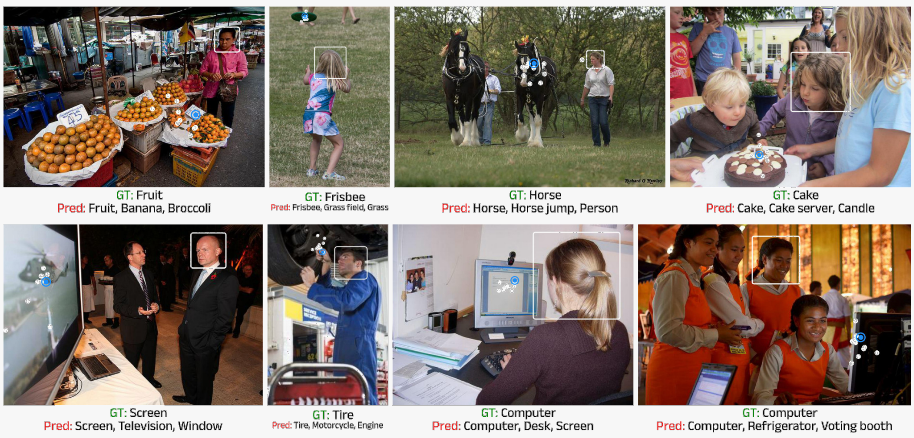
Human-AI Teaming

Semantics and social gaze attribute inference in multi-person gaze understanding

Gaze is central to human communication, revealing attention, intentions, emotions, and cognitive states. It aids in interpreting social cues, joint attention, and mental states, with applications in human-robot interaction, augmented reality (AR), education, and clinical research. Gaze Following (GF), the task of tracking where a person looks in an image or video, helps systems align with user intentions and enhance interaction modeling. However, GF is challenging as it requires analyzing both individual cues like head orientation and the broader scene context to infer targets. In [1], we tackle the prediction of gaze targets for multiple individuals in a single scene. Traditional methods process individuals sequentially, limiting real-time efficiency and accuracy. Our Sharingan model, a transformer-based architecture, predicts gaze targets simultaneously for all individuals, using as key elements gaze tokens to integrate gaze direction and spatial location. A Conditional DPT decoder (Dense Prediction Transformer) further enhances accuracy with a multiscale approach. Sharingan achieves state-of-the-art results on benchmarks, significantly improving efficiency and precision over prior methods.

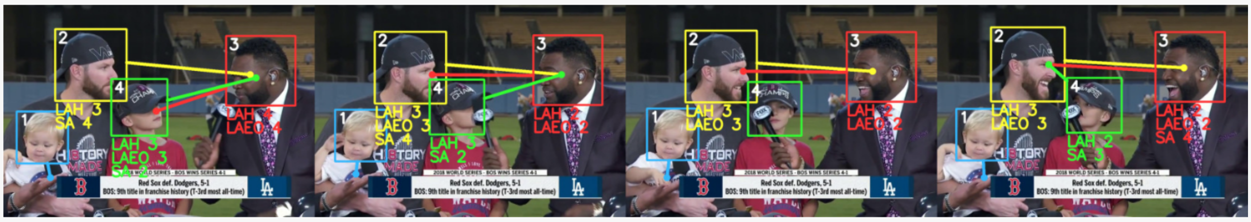
Predictions of Sharingan on naturalistic images from the internet.

In [2], we address the limitations of traditional gaze following, which predicts the locations of the gaze targets, but lacks semantic understanding. Real-world applications require recognizing not just where a person looks, but what they are looking at. Existing methods rely on separate object detection models, leading to inefficiencies, especially for ambiguous or unlabeled objects. To bridge this gap, we introduce an end-to-end framework that jointly predicts gaze target locations and semantic labels using visual-language alignment. We developed a pseudo-annotation pipeline to generate semantic labels and proposed new benchmarks for evaluation. Our approach paves the way for more holistic and meaningful applications of gaze modeling in diverse real-world settings.



Semantic predictions. Annotations are shown in white and predicted points in blue. We also provide the top three predicted class labels, and the corresponding ground-truth label below each image.

In [3], we tackle the challenge of jointly predicting gaze targets and social behaviors, which requires modeling both individual and interpersonal dynamics. Existing methods focus on either gaze following or post-processing gaze predictions, limiting accuracy and generalization, especially in multi-person, multi-frame scenarios. We propose a unified framework with a temporal transformer architecture that models interpersonal and person-scene relationships using token-based representations and cross-attention. This enables simultaneous gaze target inference and social behavior prediction, such as mutual gaze and shared attention. Backed by the VSGaze dataset, the framework achieves state-of-the-art results, showcasing the benefits of integrating gaze and social cues.



Multi-person and temporal transformer architecture for joint gaze following and social gaze prediction: Looking at Humans (LAH), Looking at Each Other (LAEO), Shared Attention (SA) [3].

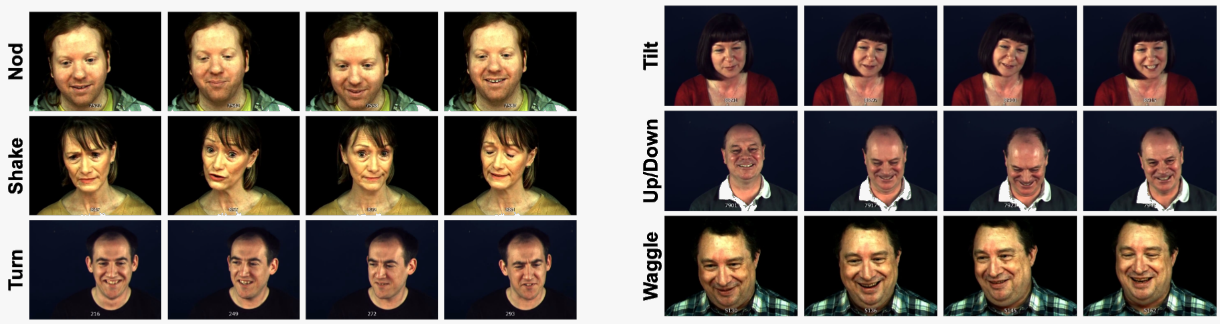
[1] S. Tafasca, A. Gupta, and J.-M. Odobez, Sharingan: A transformer architecture for multi-person gaze following, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[2] S. Tafasca, A. Gupta, V. Bros, and J.-M. Odobez, Toward semantic gaze target detection, *Advances in Neural Information Processing System (NeurIPS)*, 2024.

[3] A. Gupta, S. Tafasca, A. Farkhondeh, P. Vuillecard and J.-M. Odobez, MTGS: A novel framework for multi-person temporal gaze following and social gaze prediction, *Advances in Neural Information Processing System (NeurIPS)*, 2024.

Can gaze understanding improve the accuracy of head gesture recognition?

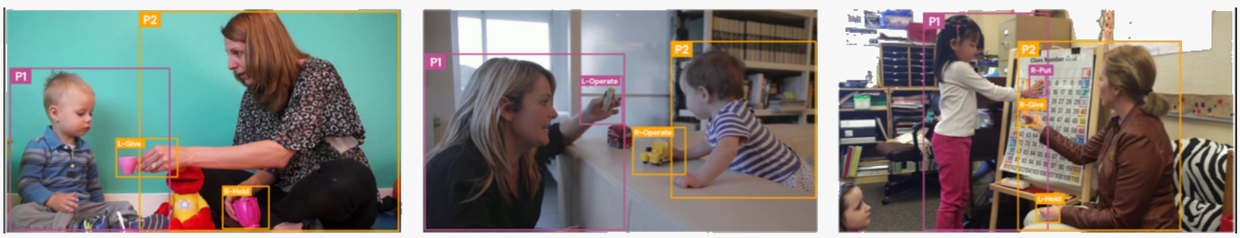
Head gestures are a vital aspect of non-verbal communication, conveying agreement, disagreement, emotions, and cognitive states in conversations. Despite their importance, existing methods for head gesture recognition (HGR) lack comprehensive datasets and often misinterpret head movements used for attention (e.g., gazing at an object) as communication gestures, leading to inaccuracies. To address this gap, we created CCDb-HG, a densely annotated dataset of diverse head gestures, incorporating gaze patterns into recognition models. By analyzing how gaze contributes to resolving ambiguities in head dynamics, we propose geometric and temporal data augmentations for better generalization across viewpoints. Our findings establish a baseline for HGR research, providing insight for using gaze to distinguish communication gestures from attention-driven head movements.



P. Vuillecard, A. Farkhondeh, M. Villamizar and J.-M. Odobez, CCDb-HG: Novel annotations and gaze-aware representations for head gesture recognition, *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, 2024.

Advancing the understanding of hand manipulations in the wild

Hand-object interactions (HOI) are pivotal in understanding human activities, with applications spanning human-computer interaction, robotics, and AR/VR. Existing datasets focus on egocentric views or controlled lab environments, real-world third-person scenarios, and often lack detailed annotations for individual hands and the inclusion of multi-person interactions. Distinguishing fine-grained hand manipulations such as grasping, holding, and operating in such dynamic scenes remains a significant challenge for action recognition and segmentation systems, especially when studying children's behaviors are underexplored. To address this gap, we introduce ChildPlay-Hand, a novel dataset featuring richly annotated videos of hand manipulations in childcare and school environments. It captures per-hand activities across various stages of manipulation. Benchmarks using state-of-the-art models highlight the dataset's complexity, emphasizing the need for hand-region-specific inputs and spatio-temporal approaches.



A. Farkhondeh, S. Tafasca, and J.-M. Odobez, Childplay-hand: A dataset of hand manipulations in the wild, *European Conference on Computer Vision (ECCV) Workshops*, 2024.

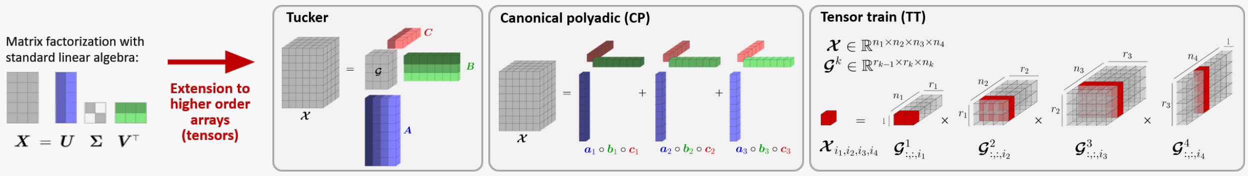
Combining end-user programming & shared autonomy for human-robot collaboration

This work addresses fundamental questions in human robot interactions: "how to specify quickly robot behaviors while ensuring their successful execution?" It brings together two human-robot collaboration approaches: end-user programming–making programming accessible to anyone regardless of their background– and shared autonomy–combining robot behaviors and real-time human corrections. This approach allows specifying complex robot behavior rapidly and adapting to specific variations of the task on-the-fly, when they are required. We evaluated this system in a sanding task, where it allows offloading the heavy manipulation to the robot, while conserving the human in the loop to ensure successful results. Insights from the development and evaluation of this system are inherently relevant to our *Human-AI Teaming* research program as it explores how humans and embodied AI (i.e., robots) can collaborate together.

M. Hagenow, E. Senft, R. Radwin, M. Gleicher, M. Zinn, and B. Mutlu, A system for human-robot teaming through end-user programming and shared autonomy, *ACM/IEEE International Conference on Human-Robot Interaction*, 2024.

Frugal robot learning with tensor networks

In contrast to standard machine learning applications, human-robot collaboration typically relies on smaller datasets, either because someone has to provide the demonstrations, or because the robot has to learn in the real physical world. We aim to design representations to efficiently transfer manipulations skills to robots by providing frugal learning strategies that can keep the most essential information from sparse data, while providing adaptation and generalization capability. To do so, we explore the use of *Tensor Networks* as a way to represent complex nonlinear functions in a separable form, with active learning strategies to estimate the models. After learning, conditional sampling is then used for fast online decision-making. In [1], we propose to use tensor trains for global optimization problems (TTGO), which has several advantages for robot skills acquisition. First, it can cope with a mix of continuous and discrete variables. Then, the active learning method used to approximate the models can easily be extended to various forms of human-guided learning, by letting the user sporadically help the robot to acquire better models (scaffolding techniques), or by letting the robot ask for help in ambiguous or unknown situations. In [2], we extended the approach to approximate reinforcement learning, with a generalized policy-iteration algorithm using tensor trains. This is an instance of approximate dynamic programming.

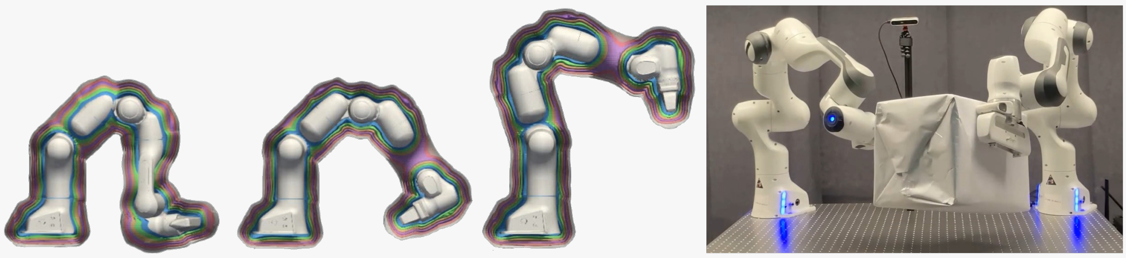


[1] S. Shetty, T. S. Lembono, T. Löw, and S. Calinon, Tensor train for global optimization problems in robotics, *International Journal of Robotics Research*, vol. 43, 2024.

[2] S. Shetty, T. Xue, and S. Calinon, Generalized policy iteration using tensor approximation for hybrid control, *International Conference on Learning Representations (ICLR)*, 2024.

Implicit shape representations for optimal control

To efficiently represent objects, tools or persons in the robot workspace, the corresponding geometric shapes can be represented in an implicit manner, by relying on *signed distance functions (SDFs)*. With this representation, the zero level set of the function corresponds to the contour of the object/shape. While many applications only focus on reconstructing this specific level set, the SDF contains additional useful information that can be leveraged in robot learning, control and planning. In [1], we developed an approach to encode SDFs as piecewise polynomial equations so that gradient and Hessian information is directly accessible in an analytic form. This differentiable representation can be used in the context of optimization and optimal control, by readily providing distance information and derivatives (gradient and Hessian for second order optimization). The gradient information additionally provides information about the normal vectors perpendicular to the surface of the objects, which is crucial for manipulation as it indicates how to orient the gripper to establish contacts with the objects to seize. In [2], we exploited these properties to build a SDF representing the whole body of a robot, with an automatic adaptation to the robot joint configuration. In [3], we showed that distance fields could also be efficiently employed in the joint angle configuration space, by providing rapid computation for whole-body manipulation, which was demonstrated in bimanual lifting of bulky objects and in interception behaviors (robotic goal keeper).



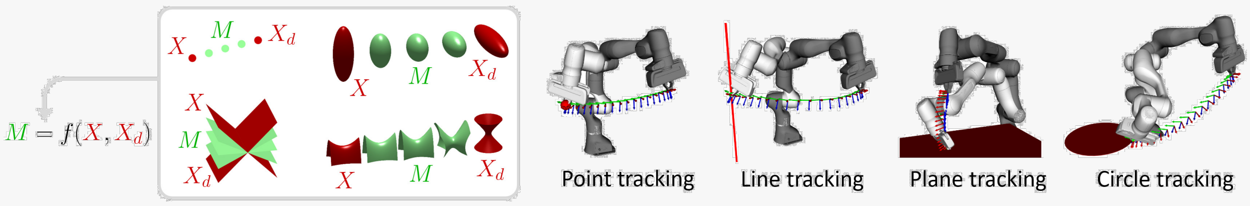
[1] A. Marić, Y. Li, and S. Calinon, Online learning of continuous signed distance fields using piecewise polynomials, *IEEE Robotics and Automation Letters (RA-L)*, vol. 9, 2024.

[2] Y. Li, Y. Zhang, A. Razmjoo, and S. Calinon, Representing robot geometry as distance fields: Applications to whole-body manipulation, *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[3] Y. Li, X. Chi, A. Razmjoo, and S. Calinon, Configuration space distance fields for manipulation planning, *Robotics: Science and Systems (RSS)*, 2024.

Geometric algebra for robotics

Because robots act in a physical world, many fundamental problems in robotics involve geometry, leading to an increased research effort in geometric methods for robotics. To speed up skills acquisition, the prior knowledge about the physical world can be embedded within the representations of skills and associated learning algorithms. This work focuses on Riemannian geometry and geometric algebra to provide such structures. Geometric algebra allows data to be processed as 32-dimensional objects describing a rich variety of geometric features, including rigid body motions and constraints specified as spheres, circles, lines or segments. It provides a generic and homogeneous formulation for various robotics problems (kinematic and dynamic control, optimization, learning and planning). Practically, it allows geometric operations to be computed in a very fast way, with compact codes [1]. This representation is thus relevant to tackle various human-robot collaboration challenges, by providing a single algebra for geometric reasoning. This work leverages the capabilities of geometric algebra in robot manipulation tasks, by modeling cost functions uniformly across different geometric primitives, with a low symbolic complexity and a geometric intuitiveness. In [1], we present a geometric algebra software library for robotics research. In [2], we apply geometric algebra to dual arm manipulators. In [3], we extend the approach to the dynamics of serial kinematic chains.



[1] T. Löw, P. Abbet, and S. Calinon, GAFRO: Geometric algebra for robotics, *IEEE Robotics and Automation Magazine*, 2024.

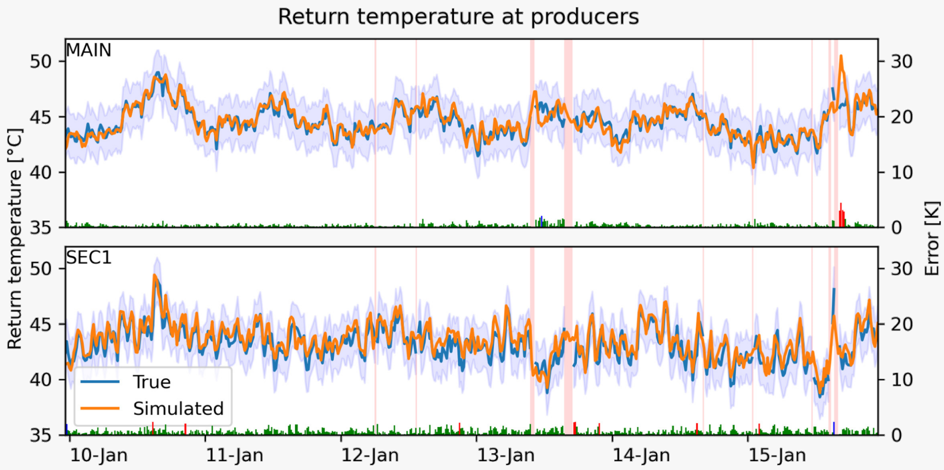
[2] T. Löw and S. Calinon, Extending the cooperative dual-task space in conformal geometric algebra, *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[3] T. Löw and S. Calinon, Recursive forward dynamics for serial kinematic chains using conformal geometric algebra, *International Workshop on the Algorithmic Foundations of Robotics (WAFR)*, 2024.

Sustainable & Resilient Societies

District heating network simulation, validation and opportunities

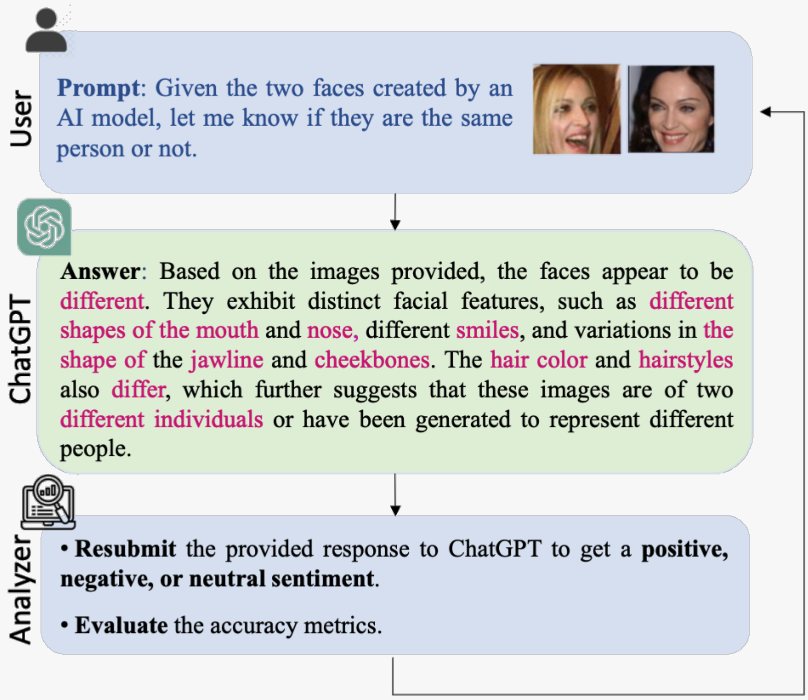
District heating networks (DHNs) play a vital role in decarbonizing the building sector, which represents one-third of global energy consumption, by enabling the recovery of waste heat. To compete economically with fossil fuel heating systems, network operators face multiple challenges: minimizing pipe losses while guaranteeing year-round supply reliability, and planning network expansions that account for future building efficiency improvements. Success requires sophisticated capabilities to dynamically optimize network control parameters and simulate network extensions based on both building demand patterns and local energy production potential. To this end, we have developed PyDHN, an open-source Python package specifically designed for third-generation DHNs. The tool implements an efficient hydraulic simulation method combined with a Lagrangian dynamic thermal model, verified against both experimental data and real network measurements from Verbier (Switzerland). Results demonstrated strong accuracy with pressure difference MAPE between 2.45-4.31% and temperature RMSE of 0.57-0.83 K. PyDHN's bottom-up approach and open-source nature make it valuable for both research applications (with reported use at the universities of Liège and Stuttgart) and practical implementations as a backend simulation engine, with future development focused on developing ML-based surrogate models.



R. Boghetti and J. H. Kämpf, Verification of an open-source Python library for the simulation of district heating networks with complex topologies, *Energy*, 2024.

ChatGPT for biometrics

We explored the application of large language models (LLMs), like ChatGPT, for biometric tasks. We specifically examine the capabilities of ChatGPT in performing biometric-related tasks, with an emphasis on face recognition, gender detection, and age estimation. Since biometrics are considered as sensitive information, ChatGPT avoids answering direct prompts, and thus we crafted a prompting strategy to bypass its safeguard and evaluate the capabilities for biometrics tasks. Our study reveals that ChatGPT recognizes facial identities and differentiates between two facial images with considerable accuracy. Additionally, experimental results demonstrate remarkable performance in gender detection and reasonable accuracy for the age estimation tasks. Our findings shed light on the promising potentials in the application of LLMs and foundation models for biometrics.



A. Hassanpour, Y. Kowsari, H. Otroshi Shahreza, B. Yang and S. Marcel, ChatGPT and biometrics: an assessment of face recognition, gender detection, and age estimation capabilities, *IEEE International Conference on Image Processing (ICIP)*, 2024.

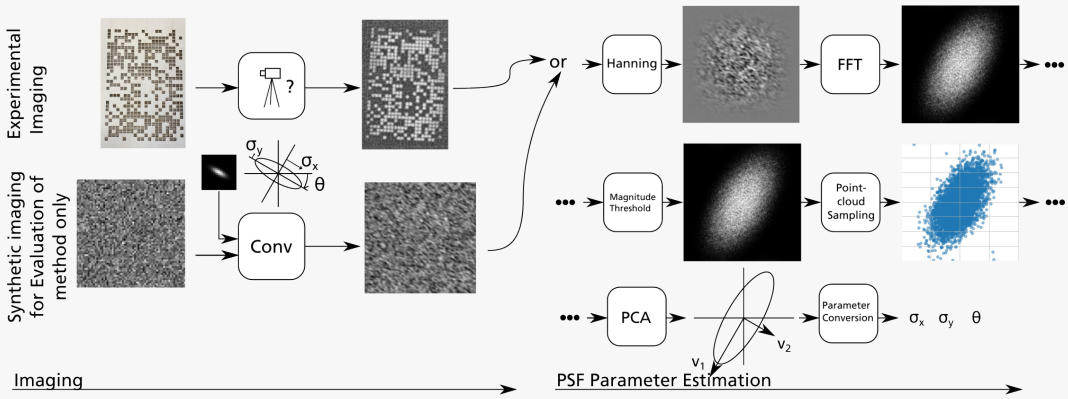
How synthetic datasets can expose real identities

Synthetic data generation is gaining increasing popularity in different computer vision applications. Existing state-of-the-art face recognition models are trained using large-scale face datasets, which are crawled from the Internet and raise privacy and ethical concerns. To address such concerns, several works have proposed generating synthetic face datasets to train face recognition models. However, these methods depend on generative models, which are trained on real face images. In this work, we design a simple yet effective membership inference attack to systematically study if any of the existing synthetic face recognition datasets leak any information from the real data used to train the generator model. We provide an extensive study on 6 state-of-the-art synthetic face recognition datasets, and show that in all these synthetic datasets, several samples from the original real dataset are leaked. To our knowledge, this paper is the first work which shows the leakage from training data of generator models into the generated synthetic face recognition datasets. Our study demonstrates privacy pitfalls in synthetic face recognition datasets and paves the way for future studies on generating responsible synthetic face datasets.

H. Otroshi Shahreza and S. Marcel, Unveiling Synthetic Faces: How Synthetic Datasets Can Expose Real Identities, *NeurIPS Workshop on New Frontiers in Adversarial Machine Learning*, 2024.

Design of imaging targets to characterize thermal cameras and imaging conditions

Imaging systems are critical for defense, surveillance, and rescue operations but are often deployed in challenging conditions that affect image quality. Our studies focused on camera motion and atmospheric turbulence. Characterizing their effect through accurate estimates of the local point spread function (PSF), which fully describes the image formation process in an image region, is an important step towards obtaining better images. We built upon techniques for estimating the PSF, using a simple image of a target object consisting of a random pattern, and described a computational pipeline for automatically estimating parametric Gaussian PSFs in the frequency domain. Through a simulation study, we compared the estimation performance when using random patterns that either followed Gaussian or binomial distributions. This allowed us to determine patterns that provide satisfactory estimation performance and that can be affordably manufactured. Based on these findings, we designed and built an experimental test target suitable for thermal imaging and experimentally demonstrated the feasibility of our approach.

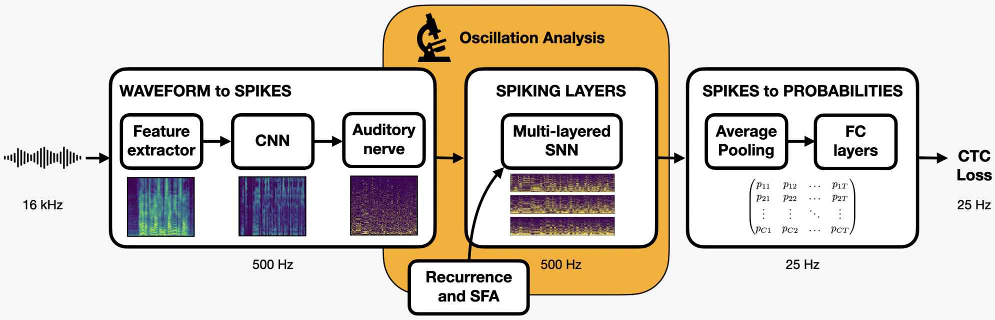


F. Piras, E. De Moura Presa, P. Wellig, and M. Liebling, Parametric point spread function estimation for thermal imaging systems using easy-to-manufacture random pattern targets, *Target and Background Signatures X: Traditional Methods and Artificial Intelligence*, vol. 13199, SPIE, 2024.

AI for Life

Artificial networks can learn biological processes

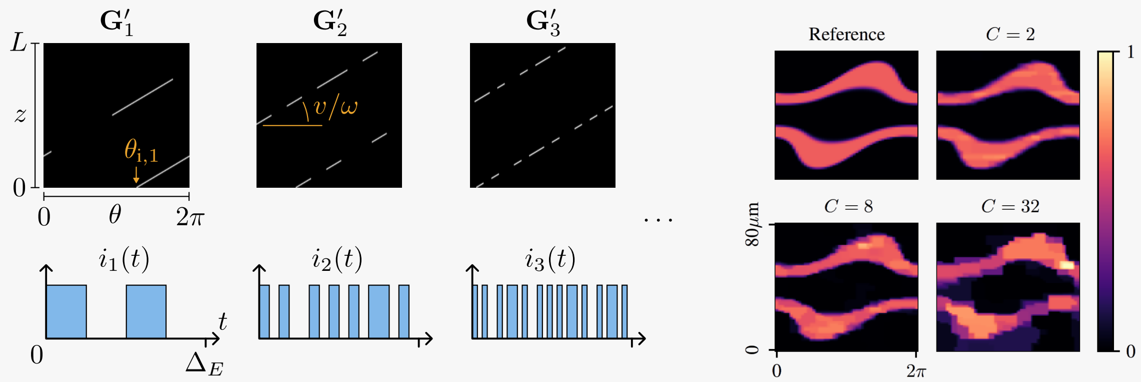
We are interested in spiking neural networks (SNNs) for two reasons. A practical reason is that they are very energy-efficient, with the potential to dramatically reduce the costs associated with training and running large networks. A second, scientific reason is that they can tell us something about the way biological processes work. If large artificial networks can do things like recognising human speech, we may hypothesise that they have learned to do so in the same manner as biological networks. In previous work, we showed that SNNs can indeed work well as speech recognisers, albeit using training techniques such as surrogate gradients that are not biologically plausible. More recently, we have been able to configure such recognisers as closely as possible to what is known about the human auditory pathway. We were able to measure oscillatory responses at each level in the SNN, characterising it in terms of frequency. We were able to observe not only the same frequencies as appear in EEG measurements, but also the same coupling (correlation) between them. The result supports the idea that we can make inferences about biological processes by building analogs that perform the same functions.



A. Bittar and P. N. Garner, Exploring neural oscillations during speech perception via surrogate gradient spiking neural networks, *Frontiers in Neuroscience*, vol. 18, 2024.

Compressive sensing for cardiac microscopy

Live microscopic imaging is an effective resource to gain understanding of heart development, both in healthy and diseased contexts. OptoMechanical Modulation Tomography (OMMT) is a compressed sensing optical microscopy method where measurements are obtained by scanning a light sheet through a sample while modulating its intensity over the course of the camera integration time. Because mechanical scanning is not instantaneous, this method was so far considered unsuitable for imaging dynamic samples. Yet, living samples would particularly benefit from the method's reduced light exposure. We have extend OMMT to allow imaging of objects that have a periodic motion, such as the heart in transparent larvae. We derived a method in which measurements are obtained by integrating the space-phase domain along patterned paths. We implemented the reconstruction with an iterative solver, and demonstrated the feasibility of the method based on simulated data of a beating heart. We observed that compression factors up to 8 lead to reliable reconstruction, and that the method is robust to uncertain acquisition start phases. Our results confirm that OMMT can be extended to imaging dynamic samples, which opens the possibility of using the method in experimental settings where low light exposure is particularly desirable.



F. Marelli and M. Liebling, Optomechanical modulation tomography for ungated compressive cardiac light sheet microscopy, *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2024.

Suppressing noise disparity in training data

Although speech-based classification approaches can achieve promising results when clean recordings are available, they may be vulnerable to additive noise. Recently, it has been shown that databases commonly used to develop and evaluate automatic pathological speech detection approaches are noisy, with the noise characteristics between healthy and pathological recordings being different. Consequently, automatic approaches trained on these databases often learn to discriminate noise rather than speech pathology. This work introduces a method to mitigate this noise disparity in training data. Using noise estimates from recordings from one group of speakers to augment recordings from the other group, the noise characteristics become consistent across all recordings. Experimental results demonstrate the efficacy of this approach in mitigating noise disparity in training data, thereby enabling automatic pathological speech detection to focus on pathology-discriminant cues rather than noise-discriminant ones.

M. Amiri and I. Kodrasi, Suppressing noise disparity in training data for automatic pathological speech detection, *International Workshop on Acoustic Signal Enhancement*, 2024.

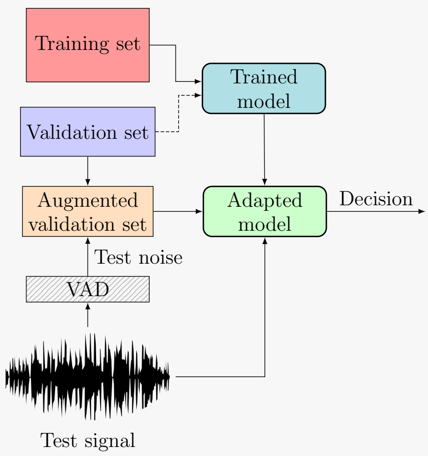
Spontaneous speech in automatic pathological speech detection

Automatic pathological speech detection approaches yield promising results in identifying various pathologies. These approaches are typically designed and evaluated for phonetically controlled speech scenarios, where speakers are prompted to articulate identical phonetic content. While gathering controlled speech recordings can be laborious, spontaneous speech can be conveniently acquired as potential patients navigate their daily routines. Furthermore, spontaneous speech can be valuable in detecting subtle and abstract cues of pathological speech. Nonetheless, the efficacy of automatic pathological speech detection for spontaneous speech remains unexplored. In this work, we analyze the influence of the speech mode on pathological speech detection approaches, examining two distinct categories of approaches, i.e., classical machine learning and deep learning. Results indicate that classical approaches may struggle to capture pathology-discriminant cues in spontaneous speech. In contrast, deep learning approaches demonstrate a higher performance, managing to extract additional cues that were previously inaccessible in controlled speech. This work paves the way towards more practical and scalable solutions, enabling real-world deployment of automatic speech detection systems that leverage everyday, naturally occurring speech to provide early and accessible diagnostic insights.

S. Shakeel and I. Kodrasi, Impact of speech mode in automatic pathological speech detection, *European Signal Processing Conference*, 2024.

Pathological speech detection adaptation in noisy environments

Deep learning-based pathological speech detection approaches are gaining popularity as a diagnostic tool to support time-consuming and subjective clinical assessments. While these approaches perform well in controlled environments with clean recordings, their performance significantly degrades in realistic scenarios with background noise. In this work, we propose a test-time adaptation framework to increase the robustness of such approaches to background noise during inference (see figure). To this end, we use a voice activity detector to extract noise-only segments from the test signal. These segments are used to augment a portion of the training/validation data, which is then exploited to fine-tune the classification models. Extensive experimental results demonstrate the effectiveness of the proposed framework in increasing robustness to noise for state-of-the-art automatic pathological speech detection approaches.



M. Amiri and I. Kodrasi, Test-time adaptation for automatic pathological speech detection in noisy environments, *European Signal Processing Conference*, 2024.

Adversarial attacks for automatic pathological speech detection

Despite the critical importance of robustness in healthcare applications like pathological speech detection, the sensitivity of deep learning-based pathological speech detection approaches to adversarial attacks remains unexplored. In this work, we explore the impact of acoustically imperceptible adversarial perturbations on pathological speech detection. Imperceptibility of perturbations, generated using the projected gradient descent algorithm, is evaluated using speech enhancement metrics. Results reveal a high vulnerability of pathological speech detection to adversarial perturbations, with adversarial training ineffective in enhancing robustness. Analyzing the perturbations provides insights into the speech components that the approaches focus on. These findings emphasize the critical need for robust automatic pathological speech detection methods suitable for clinical deployment.

M. Amiri and I. Kodrasi, Adversarial robustness analysis in automatic pathological speech detection approaches, *Annual Conference of the International Speech Communication Association*, 2024.

Parkinson's disease speech analysis and detection

As part of the SNSF Bridge Discovery project EMIL, we developed a novel syllable-based feature extraction approach for the detection of Parkinson's disease (PD). The method entails the calculation of standardised spectrotemporal patterns of syllable-like segments (fixed number of frequency and temporal bins), which are then employed as a feature vector for the detection of Parkinson's disease. We also developed a computational model based on neurophysiologically plausible computational models of speech and syllable recognition for distinguishing between healthy speech and pathological speech, and validated the proposed model on speech-based PD detection task.

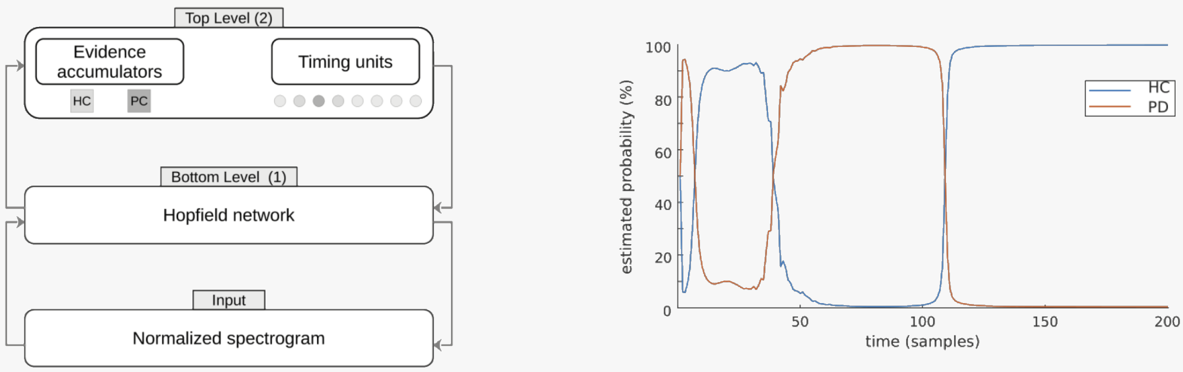


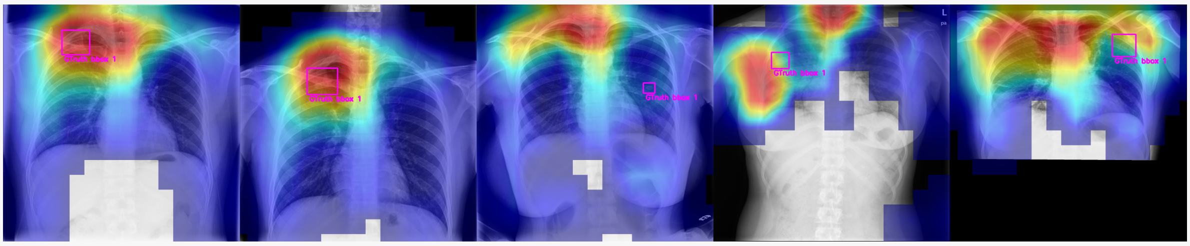
Illustration of the proposed neurocomputational model-based approach for PD detection. (Left) The top level contains evidence accumulators for PC (Parkinson condition) and HC (Healthy condition) along with timing units. (Right) The plot illustrates the dynamics of evidence accumulation. Depending on the active timing unit, the weighted spectral vector is sent to the bottom level's Hopfield network, which encodes frequency amplitude fluctuations to compare with the input spectrogram. The model follows a predictive coding framework, with arrows showing top-down predictions and bottom-up prediction errors.

S. Hovsepyan and M. Magimai-Doss, Syllable Level Features for Parkinson's Disease Detection from Speech, *IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, 2024.

S. Hovsepyan and M. Magimai-Doss, Neurocomputational model of speech recognition for pathological speech detection: a case study on Parkinson's disease speech detection, *Interspeech*, 2024.

Creating clinically interpretable AI models

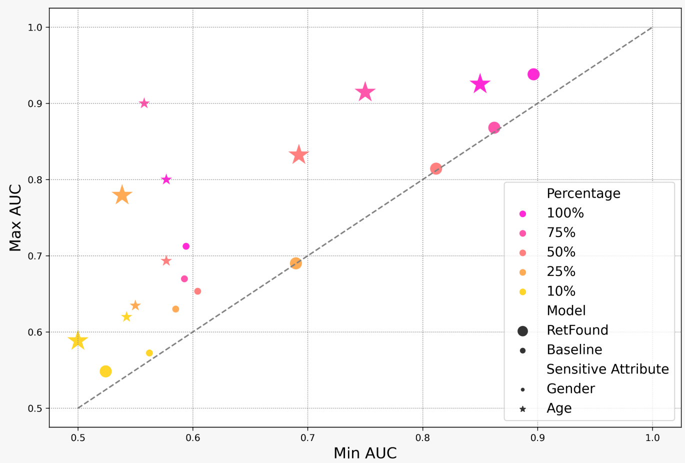
Our study addresses the critical challenge of developing reliable and interpretable AI systems for diagnosing active tuberculosis (TB) from chest X-ray (CXR) images, a vital application in resource-limited settings where expert radiologists are scarce. The research identifies that deep neural networks (DNNs), while achieving high classification accuracy, often exploit biases in datasets rather than clinically relevant features, undermining their trustworthiness. To mitigate this, we propose a pipeline involving pre-training on a large-scale proxy dataset (NIH-CXR14) with multi-objective classification tasks and fine-tuning on TB-specific data (TBX11K) using mixed objective optimization (MOON) to balance class distributions. We leverage interpretability techniques based on saliency maps (Grad-CAM, HiResCAM, and Score-CAM), to ensure the models' decision-making aligns with radiological regions of interest identified by human experts. Results (see figure) demonstrate that models pre-trained and fine-tuned with balanced data not only maintain perfect classification accuracy (AUROC=1.0) on target datasets but also generalize better to external datasets (e.g., Shenzhen dataset). Furthermore, the integration of interpretability metrics, such as proportional energy, highlights improved alignment with human judgment, paving the way for deploying AI-driven diagnostic tools in clinical and public health applications.



Ö. Güler, M. Günther, and A. Anjos, Refining tuberculosis detection in CXR imaging: Addressing bias in deep neural networks via interpretability, *European Workshop on Visual Information Processing*, 2024.

Fine-tuning foundation models without sufficient data increases demographic biases

Foundation models have emerged as robust models with label efficiency in diverse domains. In medical imaging, these models contribute to the advancement of medical diagnoses due to the difficulty in obtaining labeled data. However, it is unclear whether using a large amount of unlabeled data, biased by the presence of sensitive attributes during pre-training, influences the fairness of the model. This research examines the bias in the Foundation model (RetFound) when it is applied to fine-tune the Brazilian Multilabel Ophthalmological Dataset (BRSET), which has a different population than the pre-training dataset. The model evaluation, in comparison with supervised learning, shows that the Foundation Model has the potential to reduce the gap between the maximum AUC and minimum AUC evaluations across gender and age groups. Nevertheless, in a data-efficient generalization, the model increases the bias when the data amount decreases. These findings suggest that when deploying a Foundation Model in real-life scenarios with limited data, the possibility of fairness issues should be considered.

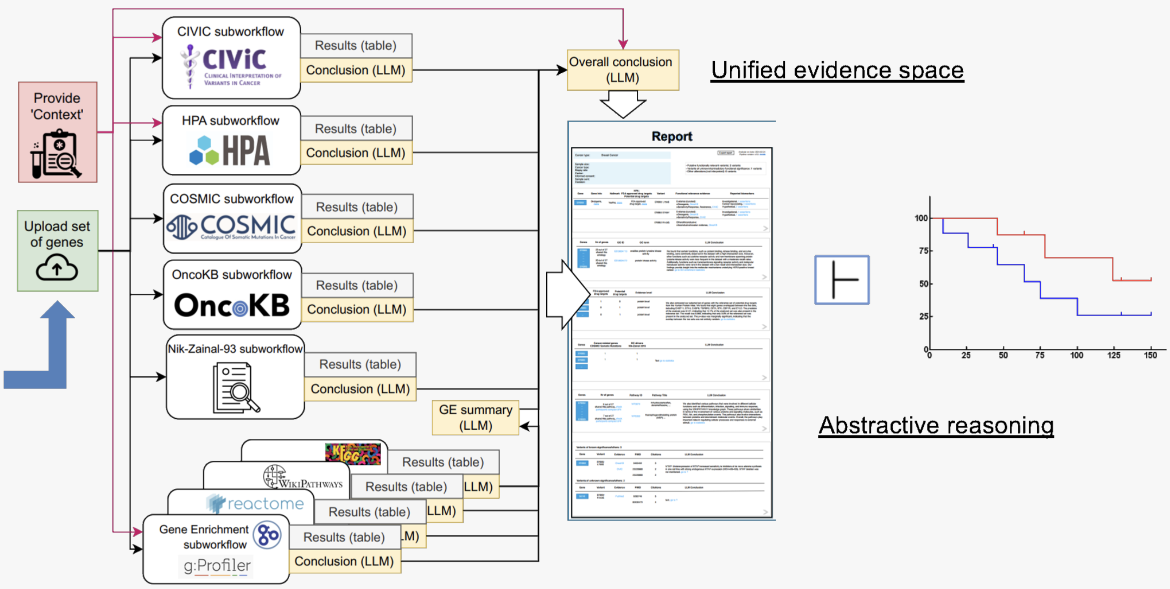


This plot shows the relationship between minimum and maximum AUC values for two models: ViT-L with RetFound weights and a baseline model developed from scratch. Point size differentiates the models, while colors and markers denote data percentages and sensitive attributes, respectively. A reference grey line enables performance comparison, succinctly encapsulating AUC dynamics and attribute interactions.

D. Queiroz Neto, A. Carlos, M. Fatoretto, L. F. Nakayama, A. Anjos, and L. Berton, Does data-efficient generalization exacerbate bias in foundation models? *European Conference on Computer Vision (ECCV)*, 2024.

Evidence-based natural language inference in oncology

Foundation models can deliver a significant value in experimental cancer medicine with a direct line-of-sight towards improving patient treatment options and outcomes. Some tumors do not respond to standard-of-care treatment options and, for some patients, emerging experimental treatments (clinical trials) offer a pathway for disease management. Due to the complexity of the disease, providing an evidence-based, personalized and precise treatment recommendation is still a significant challenge and foundation models (both language and omics-based) can provide a step-change in clinical trials management. In this context, we are extending the inference capabilities of Large Language Models (LLMs) to deliver controlled and explainable evidence-based reasoning within the Oncology domain, building neuro-symbolic models (LLMs extended with formal inference capabilities) which can support matching patients to clinical trials, reasoning over pre-clinical and clinical evidence to assist in Molecular Tumour Boards (MTBs) and to support the discovery of novel biomarkers.

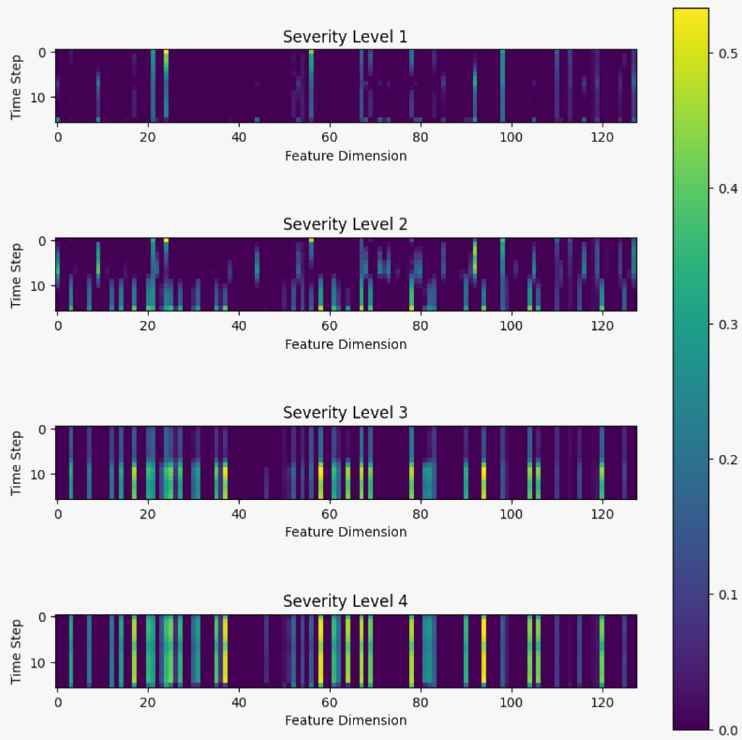


M. Jullien, M. Valentino, and A. Freitas, Semeval-2024 task 2: Safe biomedical natural language inference for clinical trials, *SemEval at NAACL*, 2024.

O. Wysocki, M. Wysocka, D. Carvalho, A. T. Bogatu, D. M. Gusicuma, M. Delmas, H. Unsworth, and A. Freitas, An llm-based knowledge synthesis and scientific reasoning framework for biomedical discovery, *Annual Meeting of the Association for Computational Linguistics (ACL), System Demonstrations Track*, 2024.

What can we learn about autism severity from long, untrimmed videos?

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition marked by unique behavioral patterns, including repetitive movements, atypical gestures, and challenges in social interactions. Early diagnosis is critical, but often delayed due to time-consuming and subjective clinical evaluations. Automating this process through technology could drastically improve the diagnosis and treatment of ASD. However, assessing ASD severity in long, untrimmed videos is complex, as it involves detecting diverse behavioral biomarkers and categorizing severity accurately from these detected patterns. Learning models for recognizing these specific markers, like facial expressions or gestures, rely on time-consuming and expensive annotations; besides, the limited set of predefined patterns lack the comprehensive analysis needed to evaluate ASD in varied real-world contexts. In this work, we propose a weakly-supervised method that uses spatio-temporal video features to detect behavioral patterns indicative of ASD. It employs a modular pipeline with a visual encoder for feature extraction, a weakly-supervised network for detecting atypical behaviors, and a regression model to estimate severity levels. By relying only on long videos, this approach reduces the dependence on dense annotations, allowing the discovery of behavioral patterns for the efficient and objective assessment of the severity of ASD, as validated on real-world clinical datasets.



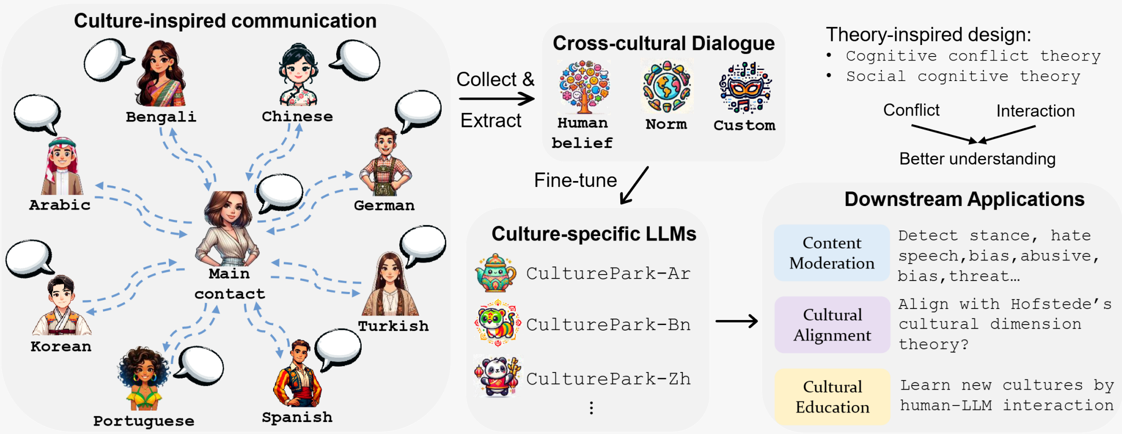
Analysis of inferred weakly-supervised temporal action localisation segments for four randomly-selected participants with four different severity levels. The patterns in the heatmaps depict various atypical biomarkers. Higher values correspond to higher severity scores.

A. Ali, M. Ali, C. Barbini, S. Dubuisson, J.-M. Odobez, F. Bremond, and S. Thümmler, Weakly-supervised autism severity assessment in long videos, *International Conference on Content-based Multimedia Indexing*, 2024.

AI for Everyone

Cultural diversity in large language models

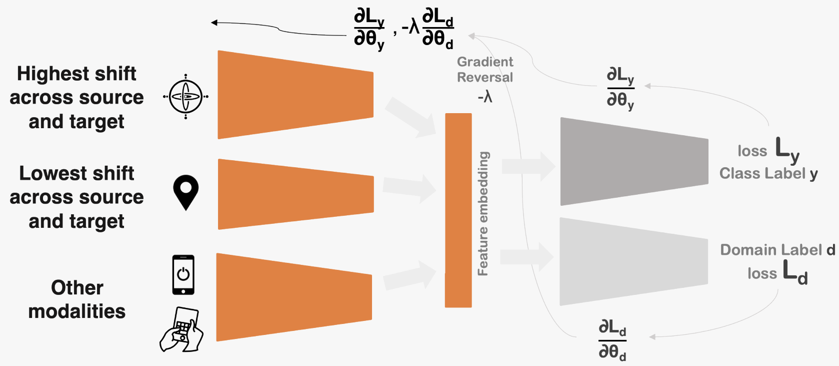
Large language models (LLMs) face a challenge: cultural bias due to being primarily trained on data from dominant cultures, which can lead to misunderstandings and misrepresentations of other cultural perspectives. In collaboration with researchers from several institutions, we developed the *CulturePark* framework to address this issue by simulating cross-cultural interactions among AI agents, each representing a different cultural background. This method generates diverse, culturally nuanced data that can be used to fine-tune LLMs, enabling them to better understand and respect a variety of cultural norms, beliefs, and values. By enhancing LLMs with a deeper sensitivity to cultural differences, CulturePark aims to support more inclusive content moderation, culturally aligned education, and broader societal value alignment. This approach aligns with the goal of making AI beneficial and adaptable to diverse communities, fostering a more culturally inclusive AI landscape.



C. Li, D. Teney, L. Yang, Q. Wen, X. Xie, and J. Wang, CulturePark: Boosting cross-cultural understanding in large language models, *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

Improving model generalization in mobile sensing

Mobile sensing has been used for inferences in contextual applications related to health and well-being. Mobile sensors can capture multimodal data, including location and accelerometer as well as device use data like application logs and screen events. A challenge limiting the deployment of these models in the real-world is the distribution shift issue, i.e., when the distribution of data in the training set differs from that of the deployment setting. This affects the generalization of models across users, populations, and environments. Using domain adversarial neural networks, we address this gap by proposing M3BAT, a technique for unsupervised domain adaptation for Multimodal Mobile sensing with Multi-branch Adversarial Training. The approach accounts for the multimodality of sensor data during domain adaptation with multiple branches. Through extensive experiments conducted on multiple mobile datasets and inference tasks, we demonstrate that our method performs effectively on unseen domains. The proposed approach shows promise towards improved model generalization, which remains a fundamental issue in the context of global populations unequally represented in terms of data availability.

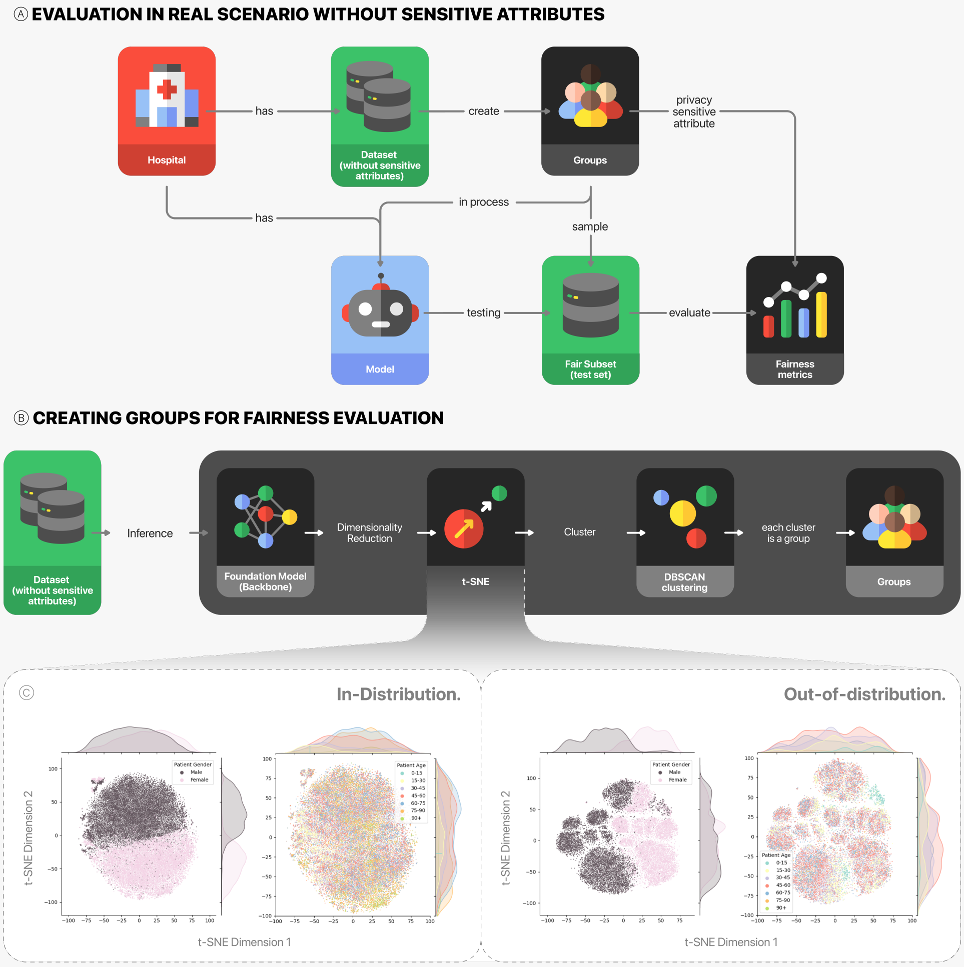


Multiple Branches handle different data modalities.

L. Meegahapola, H. Hassoune, and D. Gatica-Perez, M3BAT: Unsupervised Domain Adaptation for Multimodal Mobile Sensing with Multi-Branch Adversarial Training, *PACM on Interactive, Mobile, Wearable, and Ubiquitous Technologies, (IMWUT)*, 2024.

Improving fairness without demographic metadata

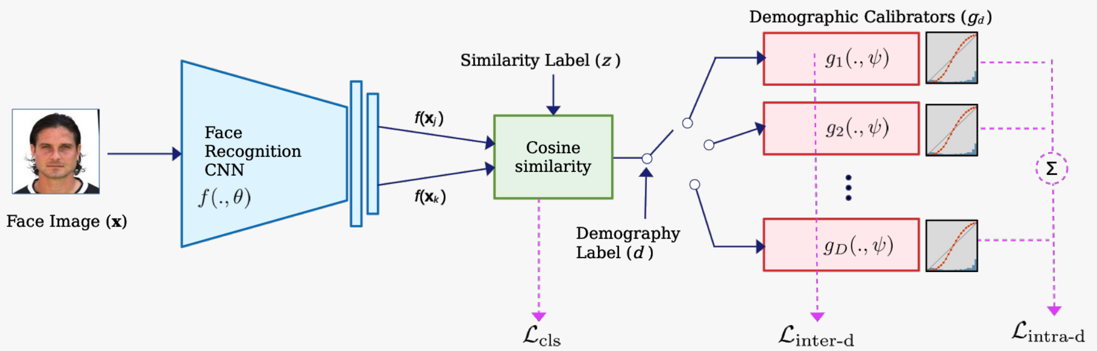
The research paper addresses the critical issue of fairness in AI-driven medical diagnostics, particularly in scenarios where demographic data—such as gender and age—are not available, which is a common challenge in medical imaging datasets. It introduces a novel approach that leverages Foundation Models (FMs), specifically their backbone, to extract meaningful embeddings from medical images and use these embeddings to infer groups that approximate sensitive attributes. These inferred groups are then applied to evaluate and mitigate biases in diagnostic AI systems. The study demonstrates the approach's effectiveness in promoting fairness for gender across in-distribution and out-of-distribution datasets, reducing disparities by up to 6.2%. However, it also highlights limitations in handling age-related attributes, signaling a need for further refinement. This methodology has significant implications for advancing equitable AI applications in healthcare by enabling fairness assessments even in the absence of explicit demographic metadata.



D. Queiroz, A. Anjos, and L. Berton, Using Backbone Foundation Model for Evaluating Fairness in Chest Radiography Without Demographic Data, *MICCAI Workshop on Fairness of AI in Medical Imaging*, 2024.

Bias mitigation in face recognition

Demographic bias in deep learning-based face recognition systems has led to serious concerns. Several existing works attempt to mitigate bias by incorporating demographic-specific processing during inference, which requires knowledge or learning of demographic attribute with an additional cost. We propose to regularize the training of the face recognition model for demographic fairness, by imposing constraints on the distributions of matching scores. Our regularization enforces score distributions from different demographic groups to respect a pre-defined distribution, and penalizes a misalignment of distributions across groups. The method improves fairness of face recognition models without compromising the recognition accuracy, and does not require extra resources during inference. Our experiments indicate that in a cross-dataset testing, the regularized CNN reduces the variation in accuracies (i.e., more fairness) across groups up to 25% while slightly improving recognition accuracy.

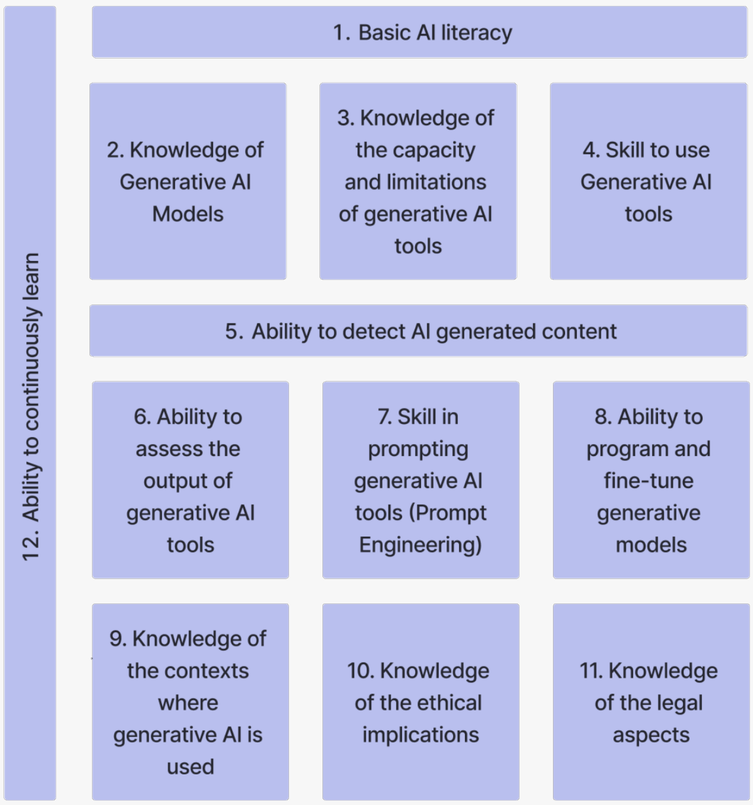


Training procedure for our regularization method for demographic fairness in face recognition. The classification loss $\mathcal{L}_{\text{cls}}$ is regularized by intra- and inter-demographic losses ($\mathcal{L}_{\text{intra-d}}$ and $\mathcal{L}_{\text{inter-d}}$). They encourage the scores from each demographic group to follow a specific distribution and align across groups.

K. Kotwal and S. Marcel, Mitigating Demographic Bias in Face Recognition via Regularized Score Calibration, *IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACV)*, 2024.

Generative AI Literacy: twelve defining competencies

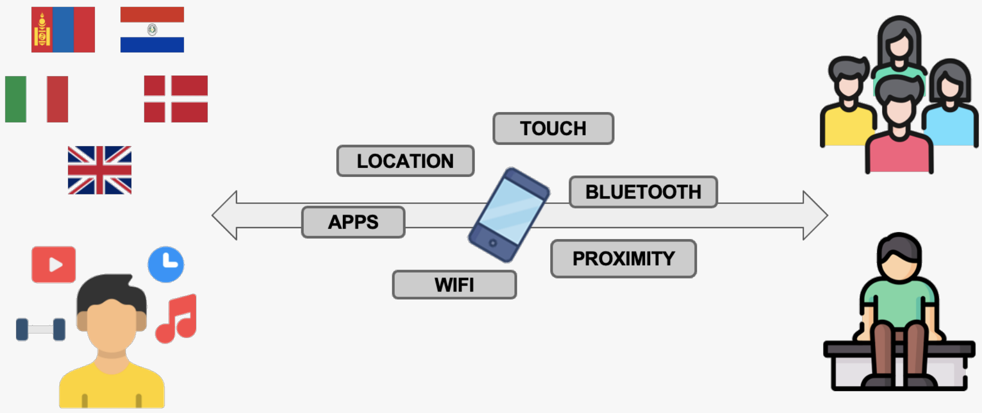
AI literacy has emerged as an important concept, not only for people specialized in technology but also for the public. The recent surge of generative AI tools has created the need for skills, abilities, and knowledge specific to GenAI's scope and applications. Understanding generative AI involves the technical aspects of the algorithms and tools, as well as contextual, ethical, and legal considerations, given the multidisciplinary potential uses and implications of these technologies. We introduced a competency-based model for GenAI literacy, defined by twelve competencies ranging from foundational AI literacy to prompt engineering and programming skills, and to ethical and legal considerations. The model can be used as a conceptual framework to assess the ability of people to understand, interact with, and create GenAI models, and can help individuals, educators, and organizations looking to navigate and take advantage of the potential of these technologies. The competencies follow a logical progression, and could serve as a starting point to develop guidelines, assessment procedures, and educational and training initiatives that support informed and responsible GenAI users and creators.



R. Annapureddy, A. Fornaroli, and D. Gatica-Perez, Generative AI Literacy: Twelve Defining Competencies, *ACM Digital Government: Research and Practice (DGOV)*, 2024.

Social context from smartphone data across countries and daily life moments

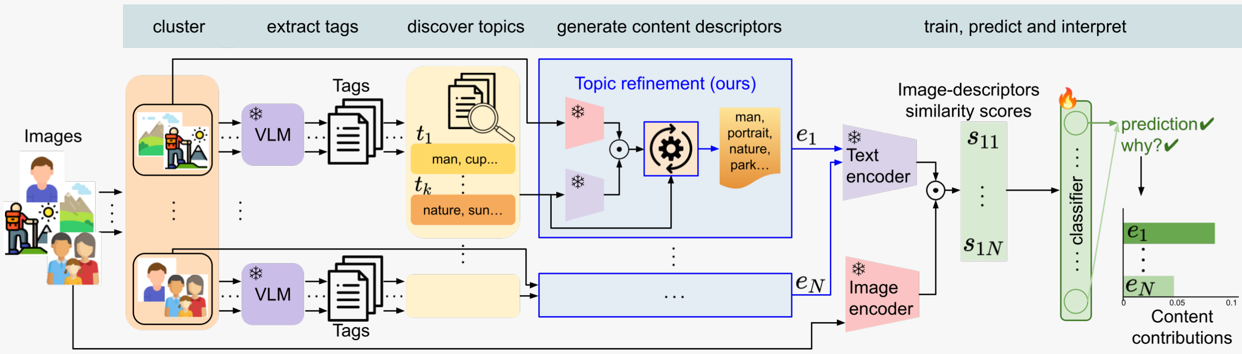
Understanding social situations in everyday life is important as part of mobile systems that support users with their personal goals and well-being. Basic social context, i.e., being alone or not, represents a starting point in this direction. The study of social context supported by mobile technology has been limited in terms of the specific daily life occasions under which social context is studied, and the geographic regions where such studies have been conducted. This in turn limits our understanding of how inference models generalize to diverse conditions, involving both daily life moments and world regions. We analyzed rich smartphone data from a cohort of 580 young adults in five countries (Mongolia, Italy, Denmark, UK, Paraguay), to understand whether this form of social context inference is feasible with smartphone data, and how country-level diversity affects inferences. We found that several sensors are informative of social context, that partially personalized multi-country models and country-specific models can achieve good performance, and that models do not generalize well to unseen countries. This highlights the importance of diversity in mobile data to build systems that work for everyone.



A. Maeder, L. Meegahapola, and D. Gatica-Perez, Learning about Social Context from Smartphone Data: Generalization Across Countries and Daily Life Moments, *ACM Conference on Human Factors in Computing Systems (CHI)*, 2024.

Interpretable image privacy classification

Images shared online can reveal personal and sensitive information that can be misused without the person's consent. How to warn users about the privacy risks associated with their images to avoid unwanted information leakage? Predicting whether an image contains sensitive information is a challenging task, especially when the decision has to be explained to a human. Post-hoc explanation methods provide only limited, hard-to-interpret insights into their decision-making. We take a human-centric approach and make the privacy decision process understandable through natural language to support people in making informed sharing decisions. Our method, called Image-guided topic modeling (ITM), enables interpretable-by-design classification based on a set of natural language content descriptors. The method leverages vision-language models (VLM) and multimodality alignment to generate the descriptors, removing the need for human-specified image attributes, which are time-consuming and costly to obtain. ITM first discovers topics from tags extracted within clusters of similar images and then refines the topics via modality alignment with the clusters' visual representation. The descriptors are then used to train linear classifiers that offer interpretability while also achieving high performance, comparable to end-to-end black-box models.



A. E. Baia and A. Cavallaro, Image-guided topic modeling for interpretable privacy classification, *European Conference on Computer Vision (ECCV) Workshops*, 2024.

Research Staff

Group Leaders

Dr. André Anjos
Medical Artificial Intelligence Group

Olivier Bornet
R&D Team

Dr. Sylvain Calinon
Robot Learning & Interaction Group

Prof. Andrea Cavallaro
Idiap Director

Dr. André Freitas
Neuro-symbolic AI Group

Dr. Philip N. Garner
Audio Inference Group

Prof. Daniel Gatica-Perez
Social Computing Group

Dr. James Henderson
Natural Language Understanding Group

Dr. Jérôme Kämpf
Energy Informatics Group

Dr. Ina Kodrasi
Signal Processing for Communication Group

Prof. Michael Liebling
Computational Bioimaging Group

Dr. Raphaëlle Luisier
Genomics & Health Informatics Group

Dr. Mathew Magimai-Doss
Speech & Audio Processing Group

Dr. Prof. Sébastien Marcel
Biometrics Security & Privacy Group

Prof. Paola Merlo
Computational Learning and Computational Linguistics Group

Dr. Petr Motlicek
Speech & Audio Processing Group

Dr. Jean-Marc Odobez
Perception & Activity Understanding Group

Dr. Emmanuel Senft
Human-centered Robotics and AI Group

Dr. Damien Teney
Machine Learning Group

Prof. Lonneke van der Plas
Computation, Cognition & Language Group

Research Associates

Name		Nationality	Supervisor	
BHATTACHARJEE	Sushil	CH	Sébastien	MARCEL
GEISSBUHLER	David	CH	Jérôme	KÄMPF
GEORGE	Anjith	IN	Sébastien	MARCEL
HOVSEPYAN	Sevada	AM	Mathew	MAGIMAI DOSS
KOMATY	Alain	FR	Sébastien	MARCEL
KORSHUNOV	Pavel	EE	Sébastien	MARCEL
KOTWAL	Ketan	IN	Sébastien	MARCEL
KRIVOKUĆA HAHN	Vedrana	HR	Sébastien	MARCEL
MADIKERI	Srikanth	IN	Petr	MOTLICEK
MOHAMMADI	Amir	IR	Sébastien	MARCEL
MURALIDHAR	Skanda	IN	Mathew	MAGIMAI DOSS
NASTASE	Viviana-Antonela	RO	Paola	MERLO
SAMO	Giuseppe	IT	Paola	MERLO
TIMONINA-FARKAS	Anna	RU	Emmanuel	SENFT
VILLAMIZAR	Michael	ES	Jean-Marc	ODOBEZ
VILLATORO TELLO	Esaú	MX	Andrea	CAVALLARO
VLASENKO	Bogdan	DE	Mathew	MAGIMAI DOSS

R&D Engineers

Name		Nationality	Supervisor	
ABBET	Philip	CH	Olivier	BORNET
BEN MAHMOUD	Imen	TN	Olivier	BORNET
CANÉVET	Olivier	CH	Olivier	BORNET
CARRON	Daniel	CH	Olivier	BORNET
CLIVAZ	Guillaume	CH	Olivier	BORNET
DAYER	Yannick	CH	Olivier	BORNET
DROZ	William	CH	Olivier	BORNET
GAIST	Samuel	CH	Olivier	BORNET
GENTILHOMME	Théophile	FR	Olivier	BORNET
KAYAL	Salim	CH	Olivier	BORNET
MACEIRAS	Jérémy	CH	Olivier	BORNET
MARCEL	Christine	CH	Olivier	BORNET
MAYORAZ	André	CH	Olivier	BORNET
MICHEL	Samuel	BE	Olivier	BORNET
NANCHEN	Alexandre	CH	Olivier	BORNET
VERZAT	Colombine	FR	Olivier	BORNET
ZANGGER	Alicia	CH	Olivier	BORNET

Post-doctoral Researchers

Name		Nationality	Supervisor	
AKSTINAITE	Vita	LT	Philip	GARNER
BAIA	Alina-Elena	RO	Andrea	CAVALLARO
BHATTACHARJEE	Mrinmoy	IN	Petr	MOTLICEK
CAROFILIS VASCO	Roberto Andrés	ES	Petr	MOTLICEK
DELMAS	Maxime	FR	André	FREITAS
ECABERT	Christophe	CH	Sébastien	MARCEL
HERMANN	Enno	DE	Mathew	MAGIMAI DOSS
HERMUS	James	US	Sylvain	CALINON
JIMÉNEZ DEL TORO	Oscar	MX	André	ANJOS
KALOGA	Yacouba	FR	Ina	KODRASI
KULKARNI	Ajinkya	IN	Mathew	MAGIMAI DOSS
LUÉVANO GARCÍA	Luis Santiago	MX	Sébastien	MARCEL
MARFURT	Andreas	CH	James	HENDERSON
MITRO	Ioanni	GR	Damien	TENEY
OTROSHI SHAHREZA	Hatef	IR	Sébastien	MARCEL
RANALDI	Leonardo	IT	André	FREITAS
RANGAPPA	Pradeep	IN	Petr	MOTLICEK
RAZEGHI	Behrooz	IR	Sébastien	MARCEL
SHEIKH	Shakeel Ahmad	IN	Ina	KODRASI
TORNAY	Sandrine	CH	Mathew	MAGIMAI DOSS
VALENTINO	Marco	IT	André	FREITAS
VÁSQUEZ RODRÍGUEZ	Laura	CR	Lonneke	VAN DER PLAS
VIDIT	Vidit	IN	Sébastien	MARCEL
WYSOCKI	Oskar	PL	André	FREITAS
XU	Zhi Ming	CA	Raphaëlle	LUISIER

PhD Students

Name		Nationality	Supervisor	
AMIRI	Mahdi	IR	Ina	KODRASI
ANNAPUREDDY	Ravinithesh Reddy	IN	Daniel	GATICA-PEREZ
BARANOUSKAYA	Darya	BY	Andrea	CAVALLARO
BEHJATI	Melika	IR	James	HENDERSON
BILALOGLU	Cem	TR	Sylvain	CALINON
BITTAR	Alexandre	CH	Philip	GARNER
BOGHETTI	Roberto	IT	Jérôme	KÄMPF
BROS	Victor	FR	Daniel	GATICA-PEREZ
CHEN	Haolin	CN	Philip	GARNER
COLBOIS	Laurent	CH	Sébastien	MARCEL
COMAN	Andrei	IT	James	HENDERSON
COPPIETERS DE GIBSON	Louise	BE	Philip	GARNER
COURDIER	Evann	FR	François	FLEURET

EL HAJAL	Karl	LB	Mathew	MAGIMAI DOSS
EL ZEIN	Dina	LB	James	HENDERSON
FARKHONDEH	Arya	IR	Jean-Marc	ODOBEZ
FEHR	Fabio	CH	James	HENDERSON
FOURNIER	Lisa	CH	Raphaëlle	LUISIER
GUPTA	Anshul	IN	Jean-Marc	ODOBEZ
HE	Mutian	CN	Philip	GARNER
HOU	Mingchi	CN	Ina	KODRASI
HRYNENKO	Olena	UA	Andrea	CAVALLARO
ISMAYILZADA	Mahammad	CA	Lonneke	VAN DER PLAS
JANKOWSKI	Julius	DE	Sylvain	CALINON
JIANG	Chunyang	CN	Paola	MERLO
JIANG	Liangze	CN	Damien	TENEY
JOHARI	Mohammad Mahdi	IR	François	FLEURET
KUMAR	Shashi	IN	Petr	MOTLICEK
LI	Yiming	CN	Sylvain	CALINON
LÖW	Tobias	CH	Sylvain	CALINON
MARIĆ	Ante	HR	Sylvain	CALINON
MATOBA	Kyle	GB	François	FLEURET
MESSORI	Elisa	IT	Raphaëlle	LUISIER
MOSTAANI	Zohreh	IR	Mathew	MAGIMAI DOSS
NIGMATULINA	Iuliia	RU	Petr	MOTLICEK
ÖZBULAK	Gökhan	TR	André	ANJOS
PETERSEN	Molly	US	Lonneke	VAN DER PLAS
PRASAD	Amrutha	IN	Petr	MOTLICEK
PUROHIT	Tilak	IN	Mathew	MAGIMAI DOSS
RAHIMI NOSHANAGH	Parsa	IR	Sébastien	MARCEL
RAJAPAKSHE	Shalutha	LK	Emmanuel	SENFT
RAZMJOO FARD	Amirreza	IR	Sylvain	CALINON
SARKAR	Eklavya	CH	Mathew	MAGIMAI DOSS
SHETTY	Suhan	IN	Sylvain	CALINON
SHIRAKAMI	Haruki	JP	James	HENDERSON
TAFASCA	Samy	MA	Jean-Marc	ODOBEZ
TARIGOPULA	Neha	IN	Mathew	MAGIMAI DOSS
UNNERVIK	Alex	CH	Sébastien	MARCEL
VUILLECARD	Pierre	CH	Jean-Marc	ODOBEZ
XUE	Teng	CN	Sylvain	CALINON
ZHANG	Yan	CN	Sylvain	CALINON
ZULUAGA GOMEZ	Juan Pablo	CO	Petr	MOTLICEK

Research Assistants

Name		Nationality	Supervisor	
ABDALLAH	Rafic	LB	Daniel	GATICA-PEREZ
AGARLA	Mirko	IT	Andrea	CAVALLARO
AL-DABBOUSSI	Mustapha	CH	Ina	KODRASI
ALONSO DEL BARRIO	David	ES	Daniel	GATICA-PEREZ
ARMANI RENZO	Matheus	BR	André	ANJOS
BI	Yuhan	CN	Daniel	GATICA-PEREZ
BRADICIC	Manuel	HR	Raphaëlle	LUISIER
BURDISSO	Sergio	AR	Petr	MOTLICEK
CHANG	Xiaoguang	CN	Damien	TENEY
CHI	Xuemin	CN	Sylvain	CALINON
COSTA BARBOSA	Lucas	BR	André	ANJOS
DAS	Pragna	IN	Sylvain	CALINON
DASTENAVAR	Atharva	FR	Emmanuel	SENFT
DE PASCALIS	Giuliano	CH	Jérôme	KÄMPF
DE WAHA-BAILLONVILLE	Gilles	BE	Jean-Marc	ODOBEZ
DUFAU	Maximilien	FR	Sylvain	CALINON
EL AYOUCH	Aliyasin	FR	Emmanuel	SENFT
FARSÄTER	Karin	SE	Jérôme	KÄMPF
FLATTOT	Vincent	FR	Sylvain	CALINON
FORNAROLI	Alessandro	IT	Daniel	GATICA-PEREZ
GANDARELA DE SOUZA	João Pedro	BR	André	FREITAS
GODI	Matthieu	CH	Sébastien	MARCEL
GÜLER	Özgür Acar	CH	André	ANJOS
HAEFLIGER	Garance	CH	Raphaëlle	LUISIER
HENRY	Oscar	FR	Raphaëlle	LUISIER
HUANG	Zhuosheng	BE	Daniel	GATICA-PEREZ
KAKAR	Ahmad Bilal	AF	Andrea	CAVALLARO
KERNEN	Marius	CH	Sébastien	MARCEL
KHALIL	Driss	MA	Petr	MOTLICEK
KOPP	Damian	CH	André	ANJOS
LJUNGBECK	Simon	SE	André	FREITAS
LORTHIOIS	Pierre	FR	Raphaëlle	LUISIER
MAETZ	Théo	FR	Sylvain	CALINON
MENENDEZ RUIZ DE AZUA	Pablo	ES	André	ANJOS
MUSCAT	Amanda	MT	Petr	MOTLICEK
NADERI	Maryam	IR	Mathew	MAGIMAI DOSS
NEUGBER	Samuel	DE	Sébastien	MARCEL
PANNATIER	Yvan	CH	André	ANJOS
PARK	Jung	KR	François	FOGLIA
PARRA ESPINOZA	Mario	MX	Daniel	GATICA-PEREZ
PARRA GALLEGO	Luis Felipe	CO	Mathew	MAGIMAI DOSS
PIAGET	Jehan	CH	Sylvain	CALINON
PIRAS	Florian	BE	Michael	LIEBLING
PIRO	Matteo	IT	Jérôme	KÄMPF

POLAC	Magdalena	PL	Sébastien	MARCEL
PULVIRENTI	Roberto	IT	André	ANJOS
QIU	Jiacheng	CN	Sylvain	CALINON
RAMIREZ CONTRERAS	Hector Manuel	MX	Daniel	GATICA-PEREZ
RODRIGUE	Dubon	FR	Jérôme	KÄMPF
RODRIGUES TORRAO	Filipa	PT	Raphaëlle	LUISIER
RUFAI	Amina	NG	Philip	GARNER
RUVOLO	Barbara	IT	Mathew	MAGIMAI DOSS
SANCHEZ LARA	Alejandra	MX	Petr	MOTLICEK
SANCHEZ-CORTES	Dairazalia	MX	Petr	MOTLICEK
SCHLÜSSEL	Timothé	CH	Sylvain	CALINON
SCHONGER	Martin	AT	Sylvain	CALINON
SELLIAH	Janani	CH	Sébastien	MARCEL
SHANGOLD	Jérôme	BE	Daniel	GATICA-PEREZ
SHAPOSHNIKOV	Roman	RU	Raphaëlle	LUISIER
SHARIF	Mohamed Amjed	JO	Jean-Marc	ODOBEZ
SYLA	Valmir	CH	Ina	KODRASI
THONET	Mateo	ES	Raphaëlle	LUISIER
ULUCAN	Ibrahim	TR	Sébastien	MARCEL
VAUTHEY	Adrien	CH	Daniel	GATICA-PEREZ
VERNHES	Albin	FR	Raphaëlle	LUISIER
VIDAL	Maxime	FR	Petr	MOTLICEK
VODOUHE	William	FR	Raphaëlle	LUISIER
WANG	Zihao	CN	Daniel	GATICA-PEREZ
XI	Lei	CN	Emmanuel	SENFT
XU	Yixuan	CN	Lonneke	VAN DER PLAS
ZATEZALO	Sara	RS	Andrea	CAVALLARO
ZENG	Kerui	CN	André	ANJOS
ZHANG	Xiaocheng	CN	Jean-Marc	ODOBEZ

Academic Visitors

Name		Nationality	Supervisor	
ALVARENGA CARDOSO	Lucas	BR	André	ANJOS
ATTINGER	Daniel	CH	Sébastien	MARCEL
BOSSHARD	Alexandra	CH	Paola	MERLO
BOU SLEIMAN	Maroun	CH	André	FREITAS
DE REZENDE ROCHA	Anderson	BR	Sébastien	MARCEL
GUSICUMA	Danilo	BR	André	FREITAS
PALLANCA	Olivier	FR	André	ANJOS
QUEIROZ NETO	Dilermando	BR	André	ANJOS
ROMAND	Bastien	FR	Emmanuel	SENFT
SHIBATA	Keigo	JP	Raphaëlle	LUISIER
SUPATTO	Willy	FR	Michael	LIEBLING
ZHANG	Alice	US	Mathew	MAGIMAI DOSS

Projects

Projects active in 2024 are grouped into three categories, namely Research Projects, Innovation Projects, and International Projects. We also list the projects accepted in 2024 and starting the following year.

Research Projects

1. Name **AI4AUTISM2** (Digital phenotyping of autism spectrum disorders in children)
   Funding SNSF Sinergia
   Coordinator University of Geneva
   Duration 2021-11-01 – 2025-10-31
   Partner(s) University of Applied Sciences and Arts of Southern Switzerland (SUPSI)

2. Name **AIML-VISIT** (Research visit at the University of Adelaide for advancing our understanding of modern AI and machine learning models)
   Funding SNSF Scientific Exchanges
   Coordinator Idiap Research Institute
   Duration 2024-12-01 – 2025-03-31
   Partner(s) Unviersity of Adelaide

3. Name **BOVINE-2** (Nonparametric Variational Representation Learning for Natural Language Understanding)
   Funding SNSF Division II
   Coordinator Idiap Research Institute
   Duration 2024-10-01 – 2028-09-30
   Partner(s) –

4. Name **CHASPEEPRO** (Characterisation of motor speech disorders and processes)
   Funding SNSF Sinergia
   Coordinator University of Geneva
   Duration 2021-12-01 – 2025-11-30
   Partner(s) Geneva University Hospitals; New Sorbonne University Paris 3

5. Name **C-LING** (Towards Creative systems with LINGuistic modelling)
   Funding SNSF Division II
   Coordinator Idiap Research Institute
   Duration 2022-09-01 – 2026-08-31
   Partner(s) –

6. Name **CODIMAN** (A future that works: Cobotics, digital skills and the re-humanization of the workplace)
   Funding SNSF NRP77
   Coordinator Berner Fachhochschule
   Duration 2020-05-01 – 2024-12-31
   Partner(s) Idiap Research Institute

7. Name **COLLABCLOUD** (A Collaborative Research Cloud Infrastructure powering Discovery and Exchange)
   Funding SNSF R'Equip
   Coordinator Idiap Research Institute
   Duration 2023-06-01 – 2024-12-31
   Partner(s) –

8. Name **DISLING** (Disentangling linguistic intelligence: automatic generalisation of structure and meaning across languages)
   Funding SNSF Advanced Grant
   Coordinator Idiap Research Institute
   Duration 2022-08-01 – 2026-07-31
   Partner(s) –

9. Name **EMIL** (Emotion in the loop – a step towards a comprehensive closed-loop deep brain stimulation in Parkinson's disease)
   Funding SNSF Bridge Discovery
   Coordinator University of Bern
   Duration 2021-05-01 – 2025-04-30
   Partner(s) Centre suisse d'électronique et de microtechnique

10. Name **EPOC** (Personalized speech recognition for audio messaging on the edge)
    Funding SNSF Birdge Proof of Concept
    Coordinator Idiap Research Institute
    Duration 2023-06-01 – 2024-05-31
    Partner(s) –

11. Name **EVOLANG** (Evolving Language Phase 1)
    Funding SNSF NCCR
    Coordinator University of Zurich
    Duration 2020-06-01 – 2024-05-31
    Partner(s) Ecole Polytechnique Federale de Lausanne; Eidgenoessische Technische Hochschule Zuerich; University of Basel; University of Fribourg; University of Geneva; University of Lausanne; University of Neuchâtel; Zurich University of the Arts

12. Name **EVOLANG-2** (Evolving Language Phase 2)
    Funding SNSF NCCR
    Coordinator University of Zurich
    Duration 2024-06-01 – 2028-05-31
    Partner(s) Ecole Polytechnique Federale de Lausanne; Eidgenoessische Technische Hochschule Zuerich; University of Basel; University of Fribourg; University of Geneva; University of Lausanne; University of Neuchâtel; Zurich University of the Arts

13. Name **EYE-TRACKING** (Creating a Multilingual Eye-Tracking Corpus for Human and Machine-Based Language Processing)
    Funding University of Zurich
    Coordinator University of Zurich
    Duration 2022-09-01 – 2025-08-31
    Partner(s) Idiap Research Institute

14. Name **FACTCHECK** (Collaborative Fact Checking of Digital Media)
    Funding Hasler Foundation
    Coordinator Idiap Research Institute
    Duration 2024-06-01 – 2025-05-31
    Partner(s) –

15. Name **FAIRMI** (Machine Learning Fairness with Application to Medical Images)
    Funding SNSF Division II
    Coordinator Federal University of Sao Paulo
    Duration 2024-03-01 – 2028-02-29
    Partner(s) Idiap Research Institute

16. Name **FORESTS** (Robust and Reliable Machine Learning for Forest Monitoring in the Valaisan Alps)
    Funding Etat du Valais
    Coordinator Idiap Research Institute
    Duration 2024-03-15 – 2025-03-31
    Partner(s) EPFL Valais Wallis

17. Name **INTREPID** (Automated interpretation of political and economic policy documents: Machine learning using semantic and syntactic information)
    Funding SNSF Sinergia
    Coordinator Graduate Institute of International and Development Studies
    Duration 2019-01-01 – 2024-03-31
    Partner(s) Idiap Research Institute

18. Name **KREATIV** (Installation at Phänomena: The roles of AI and robotics for creative and artistic applications in both digital and physical worlds)
    Funding SNSF Agora
    Coordinator Idiap Research Institute
    Duration 2024-07-01 – 2025-12-31
    Partner(s) Phänomena

19. Name **MELAS** (Alternative splicing and polyadenylation from single-cell RNA sequencing towards tumor subpopulation identification in melanoma)
    Funding Novartis Foundation
    Coordinator Idiap Research Institute
    Duration 2023-01-01 – 2024-12-31
    Partner(s) –

20. Name **NAST** (Neural Architectures for Speech Technology)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2020-02-01 – 2024-09-30
    Partner(s) –

21. Name **NEUMATH** (Neural discourse inference over mathematical texts)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2022-04-01 – 2025-03-31
    Partner(s) –

22. Name **NEUROCIRT** (Investigating the role of cytoplasmic intronic sequences in ALS pathogenesis)
    Funding SNSF Division III
    Coordinator Idiap Research Institute
    Duration 2022-09-01 – 2026-08-31
    Partner(s) –

23. Name **NEWSONAI** (Scientific voice to latest news of AI technology 2022)
    Funding SNSF Agora
    Coordinator Idiap Research Institute
    Duration 2023-04-01 – 2024-09-30
    Partner(s) –

24. Name **NKBP** (Deep learning models for continual extraction of knowledge from text)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2020-10-01 – 2025-03-31
    Partner(s) Katholieke Universiteit Leuven

25. Name **OPENBEERS** (Open-data for Building Energy Efficiency, Renovation and Storage)
    Funding Etat du Valais
    Coordinator HES-SO Valais
    Duration 2024-08-01 – 2025-07-31
    Partner(s) Idiap Research Institute

26. Name **PASS** (Pathological Speech Synthesis)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2024-03-01 – 2028-02-29
    Partner(s) –

27. Name **PAUSE** (Pathological speech enhancement)
    Funding SNSF Division II
    Coordinator Idiap Research Institute
    Duration 2024-04-01 – 2028-03-31
    Partner(s) –

28. Name **SAFER** (reSponsible fAir FacE Recognition)
    Funding Hasler Foundation
    Coordinator Idiap Research Institute
    Duration 2022-03-01 – 2025-02-28
    Partner(s) SICPA; University of Zurich

29. Name: **SMILE-II** (Scalable Multimodal sign language technology for sIgn language Learning and assessmEnt Phase-II)
    Funding: SNSF Sinergia
    Coordinator: Idiap Research Institute
    Duration: 2021-01-01 – 2024-12-31
    Partner(s): University of Applied Sciences of Special Needs Education; University of Surrey; University of Zurich

30. Name: **STEADI** (Storytelling and first impressions in face-to-face and algorithm-powered digital interviews)
    Funding: SNSF Division I
    Coordinator: University of Neuchâtel
    Duration: 2021-02-01 – 2026-05-31
    Partner(s): University of Lausanne

31. Name: **SWITCH** (Learning by switching roles in physical human-robot collaboration)
    Funding: SNSF Division II
    Coordinator: Idiap Research Institute
    Duration: 2021-03-01 – 2024-02-29
    Partner(s): Jozef Stefan Institute

32. Name: **TIPS** (Towards Integrated processing of Physiological and Speech signals)
    Funding: SNSF Division II
    Coordinator: Idiap Research Institute
    Duration: 2019-12-01 – 2024-08-31
    Partner(s): Centre suisse d'électronique et de microtechnique; Dr. Regina Maria Jankowitsch, M.A. Coaching & Moderation

Innovation Projects

1. Name: **ABROAD** (Development of an NLP tool for selecting the potential sources of novel antibiotic active against multiresistant microbes)
    Funding: The Ark
    Coordinator: Inflamalps SA
    Duration: 2022-10-01 – 2024-04-15
    Partner(s): CimArk; IHMA Europe Sàrl - BioArk; Inflamalps SA

2. Name: **BIPED** (Self-driving technology to guide blind pedestrians)
    Funding: Innosuisse
    Coordinator: Fusion Lab Technologies Sàrl
    Duration: 2023-05-01 – 2025-02-01
    Partner(s): Fusion Lab Technologies Sàrl

3. Name: **EGUZKI** (An AI simulation tool for District Heating Networks to quickly assess and predict the performance of complex meshed networks)
    Funding: SFOE
    Coordinator: RWB Valais SA
    Duration: 2020-09-01 – 2024-08-30
    Partner(s): Altis Groupe SA; Oiken SA; Satom SA

4. Name: **EPARTNERS4ALL** (Personalized and blended care solution with virtual buddy for child health)
   Funding: Innosuisse
   Coordinator: Nederlandse Organisatie voor toegepast Natuurwetenschappelijk onderzoek
   Duration: 2021-11-15 – 2024-05-15
   Partner(s): Berner Fachhochschule; Delft University of Technology; Eyeware; Interactive Robotics B.V.; Leiden University Medical Center; MedVision 360 BV; Service organization for Fire Services, Crisis Management and Public Health for the Gelderla; Therapieland; Topicus Healthcare B.V.; Xpert Health Cory Care

5. Name: **IICT** (Inclusive Information and Communication Technologies)
   Funding: Innosuisse
   Coordinator: University of Zurich
   Duration: 2022-03-14 – 2026-03-13
   Partner(s): CFS GmbH Capito; Federal Office for Civil Protection (FOCP); Federal Office for the Equality of Persons with Disabilities (OEPD); Institut Icare; Swiss Federation of the Deaf (SGB-FSS); Swiss Txt; Univ. of Applied Sciences of Special Needs Education; Zürich Versicherungs-Gesellschaft AG

6. Name: **IVECT** (Impact of greening on the energy balance and thermal comfort of buildings and districts)
   Funding: SFOE
   Coordinator: HES-SO Valais
   Duration: 2020-12-01 – 2024-12-31
   Partner(s): Centre de Recherches Energétiques et Municipales; City of Zurich; Etat du Valais; Zurich University of Applied Sciences

7. Name: **MAVERICK** (Maximum evidence platform for explainable predictions of risk)
   Funding: Innosuisse
   Coordinator: Basinghall Analytics Sàrl
   Duration: 2024-03-01 – 2026-02-28
   Partner(s): Basinghall Analytics Sàrl

8. Name: **PMPM** (Predictive Manufacturing Predictive Maintenance)
   Funding: Innosuisse
   Coordinator: HES-SO Valais
   Duration: 2024-03-01 – 2025-08-31
   Partner(s): Constellium Valais SA; Eversys SA; LYSR Sàrl

9. Name: **PRIMEAID** (Privacy-pReserving bioMetric idEntification for humAnitarian aid Distribution)
   Funding: Innosuisse
   Coordinator: Idiap Research Institute
   Duration: 2024-09-02 – 2026-03-01
   Partner(s): –

10. Name **ROSALIND** (Robust anti-fraud algorithms against malicious AI-generated face images and travel documents)
    Funding Innosuisse
    Coordinator PXL Vision AG
    Duration 2024-02-01 – 2025-07-31
    Partner(s) PXL Vision AG

11. Name **SC-IBR-1** (AI-powered directional light unit for intelligent dentist chairs)
    Funding Innosuisse
    Coordinator NVISO
    Duration 2024-02-01 – 2024-11-15
    Partner(s) NVISO SA

12. Name **SEM24** (SEM24)
    Funding Innosuisse
    Coordinator Arca24.com SA
    Duration 2023-07-01 – 2025-06-30
    Partner(s) EHL; Arca24.com SA

13. Name **SENTINEL-UK** (Securing Age Estimation and Digital IDs against Presentation and Injection Threats)
    Funding Innosuisse
    Coordinator Privately SA
    Duration 2023-11-15 – 2025-04-15
    Partner(s) Age Check Certification Scheme; Age Verification Providers Association

14. Name **SINFONIA** (Social IntelligeNce FOr strategic aNticIpAtion)
    Funding Innosuisse
    Coordinator Bloom Suisse Sàrl
    Duration 2023-04-17 – 2025-04-16
    Partner(s) Bloom Suisse Sàrl

15. Name **SMART-DISPENSING** (Programming of a complex deposit task through learning from demonstration)
    Funding Innosuisse
    Coordinator Ciposa SA
    Duration 2023-03-01 – 2026-02-28
    Partner(s) Berner Fachhochschule

16. Name **AI-SENSOR** (Sensor Fusion and Active Sensing for World-View Understanding)
    Funding Industrial
    Coordinator Idiap Research Institute
    Duration 2020-02-03 – 2024-01-31
    Partner(s) ams International AG

17. Name **CAD4IED** (Computer assisted detection and grading of inflammatory eye diseases via fluorescein angiograms and novel biomarkers)
    Funding Industrial
    Coordinator Idiap Research Institute
    Duration 2022-12-01 – 2024-05-31
    Partner(s) Fondation Asile des aveugles Hôpital Jules-Gonin; Lucerne Cantonal Hospital

18. Name **EUROCONTROL** (Integrate the Automatic Speech Recognition system with eDEP, ESCAPE and audiolan)
    Funding Industrial
    Coordinator SkySoft-ATM
    Duration 2023-02-01 – 2025-01-31
    Partner(s) Eurocontrol

19. Name **MELAS-KPI** (Towards the development of predictive biomarkers for patient stratification and immunotherapy response in cancer)
    Funding Industrial
    Coordinator Idiap Research Institute
    Duration 2023-04-01 – 2024-12-31
    Partner(s) Centre Hospitalier Universitaire Vaudois; Hôpital du Valais; Novartis Institutes for BioMedical Research

International Projects

1. Name **AI4MEDIA** (A European Excellence Centre for Media, Society and Democracy)
   Funding H2020
   Coordinator Centre for Research and Technology Hellas
   Duration 2020-09-01 – 2024-08-31
   Partner(s) Aristotle University of Thessaloniki; Athens Technology Center; Commissariat à l'Energie Atomique et aux Energies Alternatives; Consiglio nazionale delle ricerche - Istituto di Scienza e Tecnologie dell'Informazione; Côte d'Azur University; Deutsche Welle; F6S Network Limited; Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V.; Globaz SA; Grassroots Arts And Research UG (Haftungsbeschrankt); HES-SO Valais; IBM Ireland Limited; Imagga Technologies Ltd; Institut de Recherche et de Coordination Acoustique Musique; Interdigital R&D France; Joanneum Research Forschungsgesellschaft mbH; Katholieke Universiteit Leuven; Modl.ai ApS; Netherlands Institute for Sound and Vision; POLITEHNICA University of Bucharest; Queen Mary University of London; Rai Radiotelevisione Italiana Spa; Rettorato Università degli Studi di Trento; Università degli Studi di Firenze; University of Amsterdam; University of Malta; VRT (Vlaamse Radio- en Televisieomroeporganisatie)

2. Name **ALIGNAI** (value-ALIGNed socio-technical systems using large-language models (LLMs))
   Funding Horizon Europe
   Coordinator Munich University of Technology
   Duration 2024-09-01 – 2028-08-31
   Partner(s) Datacation V.V; DPG Media B.V.; Ecole Polytechnique Federale de Lausanne; ESH Médias; Fujitsu Technology Solutions GmbH; Region Hovedstaden; Technical University of Denmark; Technical University of Eindhoven; Technion Israel Institute of Technology

3. Name **CARMEN** (CarMen)
   Funding Horizon Europe
   Coordinator Yncrea Méditerranée
   Duration 2024-09-01 – 2027-08-31
   Partner(s) ARCLAN'system; B.A.I Bretagne Angleterre Irlande SA; Cabinet Louis Reynaud SAS; Hellenic Police; Hochschule Darmstadt (University of Applied Sciences H-DA); Home Office; IDEMIA Identity & Security France; Kentro Meleton Asfaletas; Ministère de l'Interieur français; Thales Dis France SAS; University of Reading; Yncrea Quest

4. Name **CRITERIA** (Comprehensive data-driven Risk and Threat Assessment Methods for the Early and Reliable Identification, Validation and Analysis of migration-related risks)
   Funding H2020
   Coordinator Gottfried Wilhelm Leibniz University Hannover
   Duration 2021-09-01 – 2024-08-31
   Partner(s) ARSIS Association for the Social Support of Youth; Centre for Research and Technology Hellas; Estonian Police and Border Guard Board; European University Cyprus; General Inspectorate of Romanian Border Police; Hensoldt Analytics (previously Sail Labs Technology AG); Knowledge and Innovation Srls (Conoscenza e Innovazione); Malta Police Force; Ministry of Interior Croatia; Rijksuniversiteit Groningen; Swedish National Police Authority; University of Malta; webLyzard technology

5. Name **ELIAS** (European Lighthouse of AI for Sustainability)
   Funding Horizon Europe
   Coordinator Rettorato Università degli Studi di Trento
   Duration 2023-09-01 – 2027-08-31
   Partner(s) Aalto University; Bitdefender SRL; Centre for Research and Technology Hellas; Copenhagen University; Czech Technical University Prague; Eberhard Karls University of Tübingen; Eidgenoessische Technische Hochschule Zuerich; Engineering - Ingegneria Informatica SPA; Fondazione Bruno Kessler; Fondazione Istituto Italiano di Tecnologia; Hasso Plattner Institute for Digital Engineering gGmbH; IBM Ireland Limited; Ideas Ncbr Sp Zoo; Institut de Recherche en Informatique et en Automatique; Institut Polytechnique de Paris; Jozef Stefan Institute; Max Planck Society for the Advancement of Sciences; Politecnico di Milano; University of Bucharest; Robert Bosch GmbH; Umea Universiteit; Universita degli Studi di Milano; Universitat de Valencia; Université de Toulouse; University of Amsterdam; University of Genoa; University of Manchester; University of Modena and Reggio Emilia; Bosch Hungary; ELLIS Alicante

6. Name **ELOQUENCE** (Multilingual and Cross-cultural interactions for context-aware, and bias-controlled dialogue systems for safety-critical applications)
   Funding Horizon Europe
   Coordinator Telefonica Investigacion Y Desarrollo SA
   Duration 2024-01-01 – 2026-12-31
   Partner(s) Barcelona Supercomputing Center; Brno University of Technology; Brunel University London; Consiglio nazionale delle ricerche - Istituto di Scienza e Tecnologie dell'Informazione; Fondazione Bruno Kessler; Grantxpert Consulting Ltd.; Omilia Ltd; Privanova SAS; Transformation Lightouse, Poslovno Svetovanje, DOO; University of Essex; University of Novi Sad; Synelixis; European University Institute; InoSens

7. Name **GRAIL-2** (Generative Range and Altitude Identity Learning)
   Funding IARPA
   Coordinator University of Southern California
   Duration 2023-05-12 – 2024-07-04
   Partner(s) –

8. Name **HURO** (Leading in the brave new world: human - robot dynamics)
   Funding Horizon Europe
   Coordinator ISM University of Management and Economics
   Duration 2024-10-01 – 2026-09-30
   Partner(s) Idiap Research Institute

9. Name **ICARUS** (Innovative AppRoach to Urban Security)
   Funding H2020
   Coordinator Forum Europeen Pour La Securite Urbaine
   Duration 2020-09-01 – 2024-08-31
   Partner(s) City of Rotterdam; City of Torino; Commune de Nice; Erasmus University Rotterdam; Ethical and Legal Plus S.L.; Eurocircle Association; Fachhochschule Salzburg GmbH; Globaz SA; Greek Center for Security Studies; Landeshauptstadt Stuttgart; Lisbon Municipal Police; Makesense; Panteion University of Social and Political Sciences; Riga Municipal Police; University of Leeds; University of Salford

10. Name **INTELLIMAN** (AI-Powered Manipulation System for Advanced Robotic Service, Manufacturing and Prosthetics)
    Funding Horizon Europe
    Coordinator Alma Mater Studiorum - Università di Bologna
    Duration 2022-09-01 – 2026-02-28
    Partner(s) Bavarian Research Alliance Gmbh; Deutsches Zentrum für Luft- und Raumfahrt e.V.; Elvez, Manufacture of cable assemblies and plastics processing D.o.o.; Eurecat Foundation; Friedrich-Alexander-Universität Erlangen-Nürnberg; National Institute for Industrial Accidents Insurance Inail; Ocado Innovation Limited; Universitat Politécnica de Cataluña; University of Campania Luigi Vanvitelli; University of Genoa; University of Zurich

11. Name **META-SPOOF** (Trustworthy AR-VR Biometric Authentication)
    Funding Meta
    Coordinator Idiap Research Institute
    Duration 2023-05-01 – 2024-12-31
    Partner(s) –

12. Name **MULTIPLEYE** (Enabling multilingual eye-tracking data collection for human and machine language processing research)
    Funding COST
    Coordinator Copenhagen University
    Duration 2022-09-28 – 2026-09-27
    Partner(s) –

13. Name **POPEYE** (robust Privacy-preserving biOmetric technologies for Passengers' identification and verification at EU external borders maximising the accuracY, reliability and throughput of the rEcognition)
    Funding Horizon Europe
    Coordinator Austrian Institute of Technology
    Duration 2024-10-01 – 2027-09-30
    Partner(s) European Association for Biometrics; General Inspectorate of Romanian Border Police; Hogskolan i Halmstad; IDEMIA Identity & Security Germany AG; Katholieke Universiteit Leuven; Ministry of Interior of the Slovak Republic; Netcompany - Intrasoft; Norwegian University of Science and Technology; Quadible Greece IKE; Universiteit Twente; Vije Universiteit Brussel

14. Name **SESTOSENSO** (Physical Cognition for Intelligent Control and Safe Human-Robot Interaction)
    Funding Horizon Europe
    Coordinator University of Genoa
    Duration 2022-10-01 – 2025-09-30
    Partner(s) Alma Mater Studiorum - Università di Bologna; Centre for Research and Technology Hellas; Centro Ricerche Fiat SCPA - Groupe Stellantis; Free University of Bolzano; Inertia Technology B.V.; Institut Franco-Allemand de Recherches de Saint-Louis; Ocado Innovation Limited; Rise Research Institutes of Sweden AB; Universidad de Zaragoza; University of Latvia, Institute Of Solid State Physics; University of Ljubljana; University of Oxford

15. Name: **SOTERIA** (uSer-friendly digiTal sEcured peRsonal data and prIvacy plAtform)
    Funding: H2020
    Coordinator: Ariadnext
    Duration: 2021-10-01 – 2024-09-30
    Partner(s): Asociatia Infocons; Audencia Business School; Autonomous University of Barcelona - CVC; Erdyn Atlantique; Institut de Recherche en Informatique et en Automatique; ipcenter.at GmbH; Katholieke Universiteit Leuven; Noria Onlus; Scytl Election Technologies, S.L.U.; Servicio Vasco De Salud Osakidetza; Stelar Security Technology Law Research UG

16. Name: **TRACY** (A big-data analyTics from base-stations Registrations And Cdrs e-evidence sYstem)
    Funding: Horizon Europe
    Coordinator: PERFORMANCE Technologies S.A.
    Duration: 2023-06-01 – 2025-05-31
    Partner(s): Cosmote Mobile Telecommunications S.A; Galati County Police Inspectorate; General Police Inspectorate of the Republic of Moldova; Greek Center for Security Studies; Hellenic Police; Hochschule Fuer den Oeffentlischen Dienst in Bayern (HföD); Time Lex CVBA

17. Name: **UNDERSPEC-ROBUSTNESS** (Addressing Underspecification for Improved Fairness and Robustness in Conversational AI)
    Funding: Amazon Research Awards
    Coordinator: Idiap Research Institute
    Duration: 2023-04-01 – 2024-03-31
    Partner(s): –

18. Name: **UNIDIVE** (Universality, diversity and idiosyncrasy in language technology)
    Funding: COST
    Coordinator: Université Paris-Saclay
    Duration: 2022-09-23 – 2026-09-22
    Partner(s): Université Paris-Saclay

Projects Awarded in 2024 and Starting in 2025

1. Name: **CERTAIN** (Certification for Ethical and Regulatory Transparency in Artificial Intelligence)
   Funding: Horizon Europe
   Coordinator: IDEMIA Identity & Security France
   Duration: 2025-01-01 – 2027-12-31
   Partner(s): EIT Digital; National Bank of Greece SA; Netcompany - Intrasoft; TARTU ULIKOOL; University of Luxembourg; Digital for Planet - d4p; DEXAI - Etica Artificiale; Charokopeio Panepistimio; Bürgerenergiegemeinschaft EMPOWER; Fachhochschule St. Pölten GmbH; Univerzitetni klinicni center Maribor; D.Tsakalidis-G.Domalis OE; Univerza V Mariboru; Red Alert LABS; N Vision systems and technologies SL; INCOM SIMVULI EPIHIRISEON EPE; Global Trust Foundation

2. Name: **GAME** (Gaze and attention modeling in Evolang)
   Funding: University of Zurich
   Coordinator: University of Zurich
   Duration: 2025-01-01 – 2025-12-31
   Partner(s): Idiap Research Institute

3. Name: **HORACE** (Human-robot collaborative manipulation by considering geometry and uncertainty)
   Funding: SNSF Division II
   Coordinator: Idiap Research Institute
   Duration: 2025-03-01 – 2029-02-28
   Partner(s): –

4. Name: **MORPHYN 2** (Modelling, Optimisation and Routing using Physiological Networks)
   Funding: SNSF Spirit
   Coordinator: Idiap Research Institute
   Duration: 2025-04-01 – 2028-03-31
   Partner(s): Mahidol University

5. Name: **ORIENTER 2** (tOwards undeRstanding and modelIng the language of mENTal health disordERs)
   Funding: SNSF Spirit
   Coordinator: Idiap Research Institute
   Duration: 2025-04-01 – 2028-03-31
   Partner(s): Mathematics Research Center; National Institute of Psychiatry

6. Name: **RATIONAL** (Robust augmented retrieval for natural language inference over transformer-based models)
   Funding: SNSF Division II
   Coordinator: State University of São Paulo
   Duration: 2025-01-01 – 2028-12-31
   Partner(s): Idiap Research Institute

7. Name — **REHABOT** (Specification and correction of physical human-robot behaviors for flexible rehabilitation robots)
   Funding — SNSF Division II
   Coordinator — Idiap Research Institute
   Duration — 2025-09-01 – 2029-08-31
   Partner(s) — Centre Neu'Rhône

8. Name — **UNIASB** (Towards a unified understanding of attention and social behaviors in the wild)
   Funding — SNSF Division II
   Coordinator — Idiap Research Institute
   Duration — 2025-04-01 – 2028-03-31
   Partner(s) — –

Code

The following code repositories were released in 2024 on Idiap's channels:
https://gitlab.idiap.ch and https://github.com/idiap.

[C1] *A novel and responsible dataset for presentation attack detection on mobile devices*, https://gitlab.idiap.ch/bob/bob.paper.ijcb2024_soteria_database.

[C2] *AgePad: Age estimation from facial images*, https://gitlab.idiap.ch/bob/paper.icassp2024.agepad.

[C3] *Agnostic features for morphing attack detection*, https://gitlab.idiap.ch/bob/bob.paper.ijcb2024_agnostic_features_mad.

[C4] *ArduinoPyTwister: Interfacing with stepper motors via arduino*, https://github.com/idiap/arduino_pytwister.

[C5] *Auto-intersphinx: Automatically generating intersphinx documentation links*, https://github.com/idiap/auto-intersphinx.

[C6] *Autocrime: Tool for processing intercepted conversations in investigations*, https://github.com/idiap/autocrime.

[C7] *Bayesian interpretation of adaptive low-rank adaptation*, https://github.com/idiap/vilora.

[C8] *Bayesian parameter-efficient fine-tuning for overcoming catastrophic forgetting*, https://github.com/idiap/bayesian-peft.

[C9] *Breaking the biometric template protection barrier*, https://gitlab.idiap.ch/bob/bob.paper.fg2024_breaking_btp.

[C10] *Can language models learn analogical reasoning?* https://github.com/idiap/analogy_learning.

[C11] *CCDbHG: Head gesture recognition demo*, https://github.com/idiap/ccdbhg-head-gesture-recognition.

[C12] *Configuration space distance fields for manipulation planning*, https://github.com/idiap/cdf.

[C13] *Configuration support for Python packages and CLIs*, https://github.com/idiap/clapper.

[C14] *Coqui - a general purpose model trainer, as flexible as it gets*, https://github.com/idiap/coqui-ai-Trainer.

[C15] *Deep variational privacy funnel: General modeling with applications in face recognition*, https://gitlab.idiap.ch/biometric/icassp2024.dvpf.

[C16] *DEFT: Demographic fairness transformer for bias mitigation in face recognition*, https://gitlab.idiap.ch/bob/bob.paper.deft_ijcb2024.

[C17] *Dialog2Flow: Convert your dialogs to flows*, https://github.com/idiap/dialog2flow.

[C18] *Ergodic sketching for robot motion planning*, https://github.com/idiap/ergodic_sketching_ros.

[C19] *Examples for the GAFRO library*, https://github.com/idiap/gafro_examples.

[C20] *Face reconstruction from facial templates by learning latent space of a generator network*, https://gitlab.idiap.ch/bob/bob.paper.neurips2023_face_ti.

[C21] *Face reconstruction from partially leaked facial embeddings*, https://gitlab.idiap.ch/bob/bob.paper.icassp2024_face_ti_partial.

[C22] *Face template inversion for biometric recognition*, https://gitlab.idiap.ch/bob/bob.paper.tbiom2024_face_ti.

[C23] *Feature-wise CCC rates for top-performing handcrafted feature representations (FR)*, https://github.com/idiap/ICASSP24_Dim_SER.

[C24] *GAFRO benchmarks for robot learning*, https://github.com/idiap/gafro_benchmarks.

[C25] *GAFRO robot descriptions for simulation*, https://github.com/idiap/gafro_robot_descriptions.

[C26] *Generating interpretations of policy announcements*, https://github.com/idiap/policy-interpretations.

[C27] *GridTK: Slurm job management for humans*, https://github.com/idiap/gridtk.

[C28] *Group membership verification in biometric recognition*, https://gitlab.idiap.ch/biometric/code.group_membership_verification.

[C29] *High-resolution coded aperture for biometric recognition*, https://gitlab.idiap.ch/bob/bob.paper.sensl2023_hires_codedaperture.

[C30] *Identifying privacy personas*, https://github.com/idiap/identifying-privacy-personas.

[C31] *Image-guided topic modeling for interpretable privacy classification*, https://github.com/idiap/itm.

[C32] *Inversion of deep facial templates using synthetic data*, https://gitlab.idiap.ch/bob/bob.paper.ijcb2023_face_ti.

[C33] *Learning robot geometry as distance fields: Applications to whole-body manipulation*, https://gitlab.idiap.ch/rli/sdf-for-robotics/rdf.

[C34] *Lensless face recognition*, https://gitlab.idiap.ch/biometric/code.face_rec_lensless.

[C35] *Linkability of multiple biometric data sources*, https://gitlab.idiap.ch/bob/bob.paper.access2024_linkability_multiple.

[C36] *Logic-LfD: Logic learning from demonstrations for multi-step manipulation tasks*, https://github.com/idiap/LogicLfD.

[C37] *Mapping the media landscape: Predicting factual reporting and political bias*, https://github.com/idiap/Factual-Reporting-and-Political-Bias-Web-Interactions.

[C38] *Model pairing for biometric recognition*, https://gitlab.idiap.ch/bob/bob.paper.neurips2024_model_pairing.

[C39] *MOE-HFR: Multi-objective evolutionary framework for robust face recognition*, https://gitlab.idiap.ch/bob/bob.paper.ijcb2024_moe_hfr.

[C40] *Morphing attack generation*, https://gitlab.idiap.ch/biometric/morphgen.

[C41] *Multi-layer self-attention, a state-of-the-art model designed for wind nowcasting tasks*, https://github.com/idiap/inference-from-real-world-sparse-measurements.

[C42] *Multitask speech recognition and speaker change detection for unknown number of speakers*, https://github.com/idiap/multitask_asr_and_scd.

[C43] *On the utility of speech and audio foundation models for marmoset call analysis*, https://github.com/idiap/speech-utility-bioacoustics.

[C44] *ppsDF: Online learning of continuous signed distance fields using piecewise polynomials*, https://gitlab.idiap.ch/rli/sdf-for-robotics/ppsdf.

[C45] *Probability-aware word-confusion-network-to-text alignment approach for intent classification*, https://github.com/idiap/Word-Confusion-Network-to-Text-Alignment.

[C46] *PyDHN: Python implementation of deep hashing networks*, https://github.com/idiap/pydhn.

[C47] *PyGAFRO: Geometric algebra for robotics in Python*, https://github.com/idiap/pygafro.

[C48] *Refining CAD for biometric recognition*, https://gitlab.idiap.ch/medai/software/paper/euvip24-refine-cad-tb.

[C49] *Reliability estimation of news media sources: Birds of a feather flock together*, https://github.com/idiap/News-Media-Reliability.

[C50] *Robust manipulation primitive learning via domain contraction*, https://github.com/idiap/robust_pl.

[C51] *ROS visualization and uRDF conversion for the GAFRO library*, https://github.com/idiap/gafro_ros.

[C52] *Sharingan: A transformer architecture for multi-person gaze following*, https://github.com/idiap/sharingan.

[C53] *Sigma-GPT: A new approach to autoregressive models*, https://github.com/idiap/sigma-gpt.

[C54] *Simple but maybe too simple config management through python data classes*, https://github.com/idiap/coqui-ai-coqpit.

[C55] *Sparse: Spiking architectures towards realistic speech encoding*, https://github.com/idiap/sparse.

[C56] *Synthetic to authentic: Code for Syn2Auth paper*, https://gitlab.idiap.ch/biometric/code.syn2auth.

[C57] *Tactile ergodic exploration for robot learning*, https://github.com/idiap/TactileErgodicExploration.

[C58] *VRBiom: Virtual reality biometrics for biometric recognition*, https://gitlab.idiap.ch/bob/bob.paper.vrbiom_pad_ijcb2024.

[C59] *Vulnerability of state-of-the-art face recognition models to template inversion attack*, https://gitlab.idiap.ch/bob/bob.paper.tifs2024_face_ti.

Datasets

The following datasets were released in 2024 by Idiap on the Zenodo platform for open science: https://zenodo.org/communities/idiap/.

[D1] D. Berling and A. Pannatier, *SkySoft ATM MALAT Vertical Rate*, https://doi.org/10.34777/9W2P-DX95.

[D2] S. Bhattacharjee, D. Geissbuhler, G. Clivaz, K. Kotwal, and S. Marcel, *CandyFV: A dataset for research on biometric recognition and presentation attack detection using vascular biometrics*, https://doi.org/10.34777/86SG-TW96.

[D3] L. Colbois, D. Geissbuhler, T. Freitas Pereira, and S. Marcel, *SynMulti-PIE: A synthetic face recognition dataset containing variations in pose, illumination and expression*, https://doi.org/10.34777/RS1Y-XZ63.

[D4] M. Delmas, M. Wysocka, and A. Freitas, *Synthetic dataset for end-to-end Relation Extraction of relationships between Organisms and Natural-Products with Mixtral-8x7B-Instruct-v0.1*, https://doi.org/10.5281/ZENODO.10849377.

[D5] M. Delmas, M. Wysocka, and A. Freitas, *Synthetic dataset for end-to-end Relation Extraction of relationships between Organisms and Natural-Products with Mixtral-8x7B-Instruct-v0.1*, https://doi.org/10.5281/ZENODO.10849378.

[D6] D. Geissbuhler, H. Otroshi Shahreza, and S. Marcel, *Synthetics-DisCo: synthetic face datasets generated via latent space exploration*, https://doi.org/10.34777/2HCM-XQ82.

[D7] A. George and S. Marcel, *Digi2Real: A synthetic face dataset containing images of 20,000 unique synthetic identities*, https://doi.org/10.34777/KDVJ-JZ94.

[D8] C. JIANG, V. Nastase, G. Samo, and P. Merlo, *BLM-AgrI-Gen: Blackbird Language Matrices Subject-Verb agreement in Italian*, https://doi.org/10.34777/07KA-PH28.

[D9] C. JIANG, V. Nastase, G. Samo, and P. Merlo, *BLM-CausI-Gen: Blackbird Language Matrices Causative/Inchoative Alternation in Italian*, https://doi.org/10.34777/2FNE-CK19.

[D10] C. JIANG, V. Nastase, G. Samo, and P. Merlo, *BLM-OdI-Gen: Blackbird Language Matrices Object-Drop Alternation in Italian*, https://doi.org/10.34777/T0NN-4J70.

[D11] P. Korshunov, A. George, G. Özbulak, and S. Marcel, *UTKPAD: Replay Attack Database for Face Age Verification*, https://doi.org/10.34777/A3AC-Y573.

[D12] K. Kotwal and S. Marcel, *Vrbiom: Virtual Reality Dataset for Biometric Applications*, https://doi.org/10.34777/PMPV-8P21.

[D13] V. Krivokuća Hahn, J. Maceiras, A. A. Komaty, P. Abbet, and S. Marcel, *iCarB-Face*, https://doi.org/10.34777/GR0N-3W62.

[D14] V. Krivokuća Hahn, J. Maceiras, A. A. Komaty, P. Abbet, and S. Marcel, *iCarB-Fingerprint*, https://doi.org/10.34777/EKTM-6M84.

[D15] V. Krivokuća Hahn, J. Maceiras, A. A. Komaty, P. Abbet, and S. Marcel, *iCarB-Voice*, https://doi.org/10.34777/JY6K-BD85.

[D16] V. Nastase and P. Merlo, *BLM-AgrE: Blackbird Language Matrices Subject-Verb agreement in English*, https://doi.org/10.34777/DVXR-K394.

[D17] V. Nastase and P. Merlo, *BLM-AgrF: Blackbird Language Matrices Subject-Verb agreement in French*, https://doi.org/10.34777/8R5S-M125.

[D18] V. Nastase and P. Merlo, *BLM-AgrI: Blackbird Language Matrices Subject-Verb agreement in Italian*, https://doi.org/10.34777/EK3E-PM24.

[D19] V. Nastase and P. Merlo, *BLM-AgrR: Blackbird Language Matrices Subject-Verb agreement in Romanian*, https://doi.org/10.34777/08GC-XQ95.

[D20] H. Otroshi Shahreza and S. Marcel, *GaFaR: Reconstructed face images from facial templates extracted from face images of the mobio dataset*, https://doi.org/10.34777/9ZSX-A621.

[D21] H. Otroshi Shahreza and S. Marcel, *SynTI: Reconstructed face images of the mobio dataset*, https://doi.org/10.34777/YVA2-8D74.

[D22] P. Rahimi Noshanagh, B. Razeghi, and S. Marcel, *RealDigiFace: The more realistic version of the digiface1m*, https://doi.org/10.34777/QX4Z-MB36.

[D23] N. Ramoly, A. A. Komaty, V. Krivokuća Hahn, L. Younes, A. Montaser Awal, and S. Marcel, *Soteria: A novel and responsible dataset for presentation attack detection on mobile devices*, https://doi.org/10.34777/9JQ3-PX34.

[D24] G. Samo and P. Merlo, *BLM-CausI: Blackbird Language Matrices Causative / Inchoative verb alternations in Italian*, https://doi.org/10.34777/D21X-D359.

[D25] G. Samo and P. Merlo, *BLM-OdI: Blackbird Language Matrices Object Drop verb alternations in Italian*, https://doi.org/10.34777/P3H6-ZY63.

[D26] S. Tafasca, A. Gupta, and J.-M. Odobez, *Sharingan data for multi-person gaze following*, https://doi.org/10.34777/WVRB-9370.

[D27] P. Vuillecard, A. Farkhondeh, M. Villamizar, and J.-M. Odobez, *Ccdb-hg: Head gesture annotation of the Cardiff Conversation Database*, https://doi.org/10.34777/WZXA-2R69.

Publications

PhD Theses

[T1] M. Behjati, Discovering meaningful units from text sequences, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T2] A. Bittar, Biologically inspired spiking neural networks for speech recognition, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T3] J. Jankowski, A stochastic approach to contact-rich manipulation, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T4] M. M. Johari, Advancing self-supervised deep learning for 3d scene understanding, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T5] Z.-G. Juan, Low-resource speech recognition and understanding for challenging applications, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T6] K. Matoba, Safe deep neural networks, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T7] H. Otroshi Shahreza, On the information in deep biometric templates: From vulnerability of unprotected templates to leakage in protected templates, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T8] A. Pannatier, Extending capabilities of attention-based models, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T9] S. Shetty, Robot learning using tensor networks, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

[T10] A. Unnervik, Performing and detecting backdoor attacks on face recognition algorithms, PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2024.

Journal Papers

[J1] M. Agarwal, P. Cameron-Rastogi, G. Peronato, and G. Mavromatidis, Missed opportunities in building energy performance assessment, *Journal of Sustainable Real Estate*, vol. 16, 2024.

[J2] R. Annapureddy, A. Fornaroli, and D. Gatica-Perez, Generative AI literacy: Twelve defining competencies, *ACM Digital Government: Research and Practice*, 2024.

[J3] A. Bittar and P. N. Garner, Exploring neural oscillations during speech perception via surrogate gradient spiking neural networks, *Frontiers in Neuroscience*, vol. 18, 2024.

[J4] R. Boghetti and J. H. Kämpf, Verification of an open-source python library for the simulation of district heating networks with complex topologies, *Energy*, vol. 290, no. C, 2024.

[J5] A. B. Bosshard, J. Burkart, P. Merlo, C. Cathcart, B. Bickel, and S. Towsend, Beyond bigrams: Call sequencing in the common marmoset (callithrix jacchus) vocal system, *Royal Society Open Science*, vol. 11, 2024.

[J6] G. Braglia, S. Calinon, and L. Biagiotti, A minimum-jerk approach to handle singularities in virtual fixtures, *IEEE Robotics and Automation Letters (RA-L)*, vol. 9, 2024.

[J7] H. Chen and P. N. Garner, Bayesian parameter-efficient fine-tuning for overcoming catastrophic forgetting, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2024.

[J8] G. Degano, P. Donhauser, L. Gwilliams, P. Merlo, and N. Golestani, Speech prosody enhances the neural processing of syntax, *Communications Biology*, vol. 7, 2024.

[J9] A. George, C. Ecabert, H. Otroshi Shahreza, K. Kotwal, and S. Marcel, Edgeface : Efficient face recognition model for edge devices, *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2024.

[J10] A. George and S. Marcel, From modalities to styles: Rethinking the domain gap in heterogeneous face recognition, *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2024.

[J11] L. Coppieters de Gibson and P. N. Garner, Training a filter-based model of the cochlea in the context of pre-trained acoustic models, *Acoustics*, vol. 6, 2024.

[J12] F. Holzknecht, S. Tornay, A. Battisti, A. O. Batty, K. Tissi, T. Haug, and S. Ebling, Automated sign language vocabulary assessment: Comparing human and machine ratings and studying learner perceptions, *Language Assessment Quaterly*, 2024.

[J13] J. Huang, Z. Li, S. Bhattacharjee, W. Kang, and S. Marcel, Mirror-based full-view finger vein authentication with illumination adaptation, *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[J14] E. Huber, S. Sauppe, A. Isasi-Isasmendi, I. Bornkessel-Schlesewsky, P. Merlo, and B. Bickel, Surprisal from language models can predict ERPs in processing predicate-argument structures only if enriched by an agent preference principle, *Neurobiology of Language*, vol. 5, 2024.

[J15] Z. Liu, W. Sun, D. Teney, and S. Gould, Candidate set re-ranking for composed image retrieval with dual multi-modal encoder, *Transactions on Machine Learning Research (TMLR)*, 2024.

[J16] T. Löw, P. Abbet, and S. Calinon, GAFRO: Geometric algebra for robotics, *IEEE Robotics and Automation Magazine*, 2024.

[J17] A. Marić, Y. Li, and S. Calinon, Online learning of continuous signed distance fields using piecewise polynomials, *IEEE Robotics and Automation Letters (RA-L)*, vol. 9, 2024.

[J18] T. Mautuit, P. Cunnac, F. Truffer, A. Anjos, R. Dufrane, G. Maître, M. Geiser, and C. Chiquet, Absolute retinal blood flow in healthy eyes and in eyes with retinal vein occlusion, *Microvascular Research*, vol. 152, 2024.

[J19] L. B. Meegahapola, H. Hassoune, and D. Gatica-Perez, M3bat: Unsupervised domain adaptation for multimodal mobile sensing with multi-branch adversarial training, *PACM on Interactive, Mobile, Wearable, and Ubiquitous Technologies (IMWUT)*, vol. 8, 2024.

[J20] M. Mühlbauer, T. Hulin, B. Weber, S. Calinon, F. Stulp, A. Albu-Schäffer, and J. Silverio, A probabilistic approach to multi-modal adaptive virtual fixtures, *IEEE Robotics and Automation Letters (RA-L)*, 2024.

[J21] H. Otroshi Shahreza, V. Krivokuca, and S. Marcel, Vulnerability of state-of-the-art face recognition models to template inversion attack, *IEEE Transactions on Information Forensics and Security*, vol. 19, 2024.

[J22] H. Otroshi Shahreza and S. Marcel, Template inversion attack using synthetic face images against real face recognition systems, *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2024.

[J23] H. Otroshi Shahreza, Y. Y. Shkel, and S. Marcel, On measuring linkability of multiple protected biometric templates using maximal leakage, *IEEE Access*, 2024.

[J24] A. Pannatier, K. Matoba, and F. Fleuret, Inference from real-world sparse measurements, *Transactions on Machine Learning Research (TMLR)*, 2024.

[J25] B. Razeghi, M. Gheisari, A. Atashin, D. Kostadinov, S. Marcel, D. Gunduz, and S. Voloshynovskiy, Group membership verification via nonlinear sparsifying transform learning, *IEEE Access*, vol. 12, 2024.

[J26] S. Shetty, T. S. Lembono, T. Löw, and S. Calinon, Tensor train for global optimization problems in robotics, *International Journal of Robotics Research*, vol. 43, 2024.

[J27] I. Shiri, B. Razeghi, S. Ferdowsi, Y. Salimi, D. Gunduz, D. Teodoro, S. Voloshynovskiy, and H. Zaidi, Primis: Privacy-preserving medical image sharing via deep sparsifying transform learning with obfuscation, *Journal of Biomedical Informatics, Elsevier*, vol. 150, 2024.

[J28] B. Ti, Y. Gao, J. Zhao, and S. Calinon, An optimal control formulation of tool affordance applied to impact tasks, *IEEE Transactions on Robotics (T-RO)*, 2024.

[J29] L. Tian, C. Oh, and A. Cavallaro, Test-time adaptation for 6D pose tracking, *Pattern Recognition*, vol. 152, 2024.

[J30] B. Tobler-Ammann, F. Schuind, L. Voillat, T. Gentilhomme, E. Vögelin, N. Murith, and B. Masserey, Developing 3d-printed wrist splints for distal radius and scaphoid fractures, *Journal of Wrist Surgery*, 2024.

[J31] A. Unnervik, A. George, H. Otroshi Shahreza, P. Rahimi, and C. Ecabert, FRCSyn-onGoing: Benchmarking and comprehensive evaluation of real and synthetic data to improve face recognition systems, *Information Fusion*, vol. 107, 2024.

[J32] C. Volf, B. Bueno, P. Edwards, R. Hobday, S. Mäder, B. Matusiak, K. Wulff, W. Osterhaus, G. Manoli, C. Della Giustina, J. Joshi, J. Kämpf, K. Vega, and C. Kueffer, Why daylight should be a priority for urban planning, *Journal of Urban Management*, 2024.

[J33] J. Wang, S. Kim, T. S. Lembono, W. Du, J. Shim, S. Samadi, K. Wang, V. Ivan, S. Calinon, S. Vijayakumar, and S. Tonneau, Online multi-contact receding horizon planning via value function approximation, *IEEE Transactions on Robotics (T-RO)*, 2024.

[J34] Y. Zhang, T. Xue, A. Razmjoo, and S. Calinon, Logic learning from demonstrations for multi-step manipulation tasks in dynamic environments, *IEEE Robotics and Automation Letters*, 2024.

Conference Papers

[1] M. Agarla and A. Cavallaro, GLoFool: Global enhancements and local perturbations to craft adversarial images, *European Conference on Computer Vision (ECCV) Workshops*, 2024.

[2] Z. Ahmadi, Z. Zhang, H. H. Nguyen, S. Burdisso, S. Madikeri, P. Motlicek, E. Dikici, G. Backfried, M. Kovac, and D. Kudenko, Entity matching across small networks using node attributes, *European Conference on Artificial Intelligence*, 2024.

[3] A. Ali, M. Ali, C. Barbini, S. Dubuisson, J.-M. Odobez, F. Bremond, and S. Thümmler, Weakly-supervised autism severity assessment in long videos, *International Conference on Content-based Multimedia Indexing*, 2024.

[4] M. Amiri and I. Kodrasi, Adversarial robustness analysis in automatic pathological speech detection approaches, *Annual Conference of the International Speech Communication Association*, 2024.

[5] M. Amiri and I. Kodrasi, Suppressing noise disparity in training data for automatic pathological speech detection, *International Workshop on Acoustic Signal Enhancement*, 2024.

[6] M. Amiri and I. Kodrasi, Test-time adaptation for automatic pathological speech detection in noisy environments, *European Signal Processing Conference*, 2024.

[7] T. Apicella, A. Xompero, P. Gastaldo, and A. Cavallaro, Segmenting object affordances: Reproducibility and sensitivity to scale, *European Conference on Computer Vision (ECCV) Workshops*, 2024.

[8] A. E. Baia and A. Cavallaro, Image-guided topic modeling for interpretable privacy classification, *European Conference on Computer Vision (ECCV) Workshops*, 2024.

[9] D. Alonso del Barrio, M. Tiel, and D. Gatica-Perez, Human interest or conflict? leveraging LLMs for automated framing analysis in tv shows, *ACM International Conference on Interactive Media Experiences*, 2024.

[10] I. Ben Mahmoud, E. Sarkar, M. Manser, and M. Magimai-Doss, Feature representations for automatic meerkat vocalization classification, *International Workshop on Vocal Interactivity in-and-between Humans, Animals and Robots*, 2024.

[11] M. Bhattacharjee, N. Iuliia, A. Prasad, P. Rangappa, S. Madikeri, P. Motlicek, H. Helmke, and M. Kleinert, Contextual biasing methods for improving rare word detection in automatic speech recognition, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[12] S. Bhattacharjee, D. Geissbuhler, G. Clivaz, K. Kotwal, and S. Marcel, Vascular biometrics experiments on candy – a new contactless finger-vein dataset, *International Conference on Pattern Recognition (ICPR)*, 2024.

[13] S. Burdisso, S. Madikeri, and P. Motlicek, Dialog2flow: Pre-training soft-contrastive action-driven sentence embeddings for automatic dialog flow extraction, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, ACL, 2024.

[14] S. Burdisso, E. A. Reyes-Ramírez, E. Villatoro-Tello, F. Sánchez-Vega, A. P. López-Monroy, and P. Motlicek, Daic-woz: On the validity of using the therapist's prompts in automatic depression detection from clinical interviews, *Clinical Natural Language Processing Workshop*, Association for Computational Linguistics, 2024.

[15] S. Burdisso, D. Sanchez-Cortes, E. Villatoro-Tello, and P. Motlicek, Reliability estimation of news media sources: Birds of a feather flock together, *Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.

[16] L. Colbois and S. Marcel, Evaluating the effectiveness of attack-agnostic features for morphing attack detection, *International Joint Conference on Biometrics*, 2024.

[17] A. C. Coman, C. Theodoropoulos, M.-F. Moens, and J. Henderson, GADePo: Graph-assisted declarative pooling transformers for document-level relation extraction, *Workshop on Knowledge Augmented Methods for NLP*, Association for Computational Linguistics, 2024.

[18] J. Corsetti, D. Boscaini, C. Oh, A. Cavallaro, and F. Poiesi, Open-vocabulary object 6D pose estimation, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[19] G. Elbanna, Z. Mostaani, and M. Magimai-Doss, Predicting heart activity from speech using data-driven and knowledge-based features, *Annual Conference of the International Speech Communication Association (Interspeech)*, 2024.

[20] A. Espuña, A. Prasad, P. Motlicek, S. Madikeri, and S. Christof, Normalizing flows for speaker and language recognition backend, *Odyssey 2024: The Speaker and Language Recognition Workshop*, 2024.

[21] A. Farkhondeh, S. Tafasca, and J.-M. Odobez, Childplay-hand: A dataset of hand manipulations in the wild, *European Conference on Computer Vision (ECCV) Workshops*, 2024.

[22] F. Fehr and J. Henderson, Nonparametric variational regularisation of pretrained transformers, *Conference on Language Modelling (COLM)*, 2024.

[23] S. Gao, M. Ismayilzada, M. Zhao, H. Wakaki, Y. Mitsufuji, and A. Bosselut, DiffuCOMET: Contextual commonsense knowledge diffusion, *Annual Meeting of the Association for Computational Linguistics (ACL)*, L.-W. Ku, A. Martins, and V. Srikumar, Eds., ser. Long Papers, vol. 1, Association for Computational Linguistics, 2024.

[24] A. George and S. Marcel, Heterogeneous face recognition using domain invariant units, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[25] A. George and S. Marcel, Modality agnostic heterogeneous face recognition with switch style modulators, *IEEE International Joint Conference on Biometrics*, 2024.

[26] Ö. Güler, M. Günther, and A. Anjos, Refining tuberculosis detection in CXR imaging: Addressing bias in deep neural networks via interpretability, *European Workshop on Visual Information Processing*, 2024.

[27] A. Gupta, S. Tafasca, N. Chutisilp, and J.-M. Odobez, A unified model for gaze following and social gaze prediction, *IEEE International Conference on Automatic Face and Gesture Recognition*, 2024.

[28] A. Gupta, S. Tafasca, A. Farkhondeh, P. Vuillecard, and J.-M. Odobez, Mtgs: A novel framework for multi-person temporal gaze following and social gaze prediction, *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

[29] A. Gupta, P. Vuillecard, A. Farkhondeh, and J.-M. Odobez, Exploring the zero-shot capabilities of vision-language models for improving gaze following, *Workshop on Gaze Estimation and Prediction in the Wild at CVPR*, 2024.

[30] M. Hagenow, E. Senft, R. Radwin, M. Gleicher, M. Zinn, and B. Mutlu, A system for human-robot teaming through end-user programming and shared autonomy, *ACM/IEEE International Conference on Human-Robot Interaction*, 2024.

[31] A. Hassanpour, Y. Kowsari, H. Otroshi Shahreza, B. Yang, and S. Marcel, ChatGPT and biometrics: An assessment of face recognition, gender detection, and age estimation capabilities, *IEEE International Conference on Image Processing (ICIP)*, 2024.

[32] S. Hovsepyan and M. Magimai-Doss, Neurocomputational model of speech recognition for pathological speech detection: A case study on parkinson's disease speech detection, *Annual Conference of the International Speech Communication Association (Interspeech)*, 2024.

[33] S. Hovsepyan and M. Magimai-Doss, Parkinson's disease detection through formant and F0 analysis at syllable level, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[34] S. Hovsepyan and M. Magimai-Doss, Syllable level features for parkinson's disease detection from speech, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[35] L. Igene, A. Hossain, M. Z. U. Chowdhury, H. Rezaie, A. Rollins, J. Dykes, R. Vijaykumar, A. Komaty, S. Marcel, S. Schuckers, J. E. Tapia, C. Aravena, D. Schulz, B. Adami, N. Karimian, D. Nunes, J. Marcos, N. Gonçalves, L. Sikosek, B. Batagelj, A. Alenin, A. Alkhaddour, A. Pimenov, A. Tregubov, I. Avdonin, M. Kazantsev, M. Pozigun, V. Pryadchenko, N. Schei, D. Pabon, and M. Tiedemann, Face liveness detection competition (LivDet-Face), *IEEE International Joint Conference on Biometrics*, 2024.

[36] C. Jiang, G. Samo, V. Nastase, and P. Merlo, Blm-it - blackbird language matrices for italian: A calamita challenge, *Italian Conference on Computational Linguistics*, 2024.

[37] L. Jiang and D. Teney, Can we learn to select the right algorithm for ood generalization? *Out Of Distribution Generalization in Computer Vision, Workshop at ECCV*, 2024.

[38] M. Jullien, M. Valentino, and A. Freitas, Semeval-2024 task 2: Safe biomedical natural language inference for clinical trials, *SemEval at NAACL*, 2024.

[39] V. Jung and L. van der Plas, Understanding the effects of language-specific class imbalance in multilingual fine-tuning, *Conference of the European chapter of Association for Computational Linguistics (EACL)*, 2024.

[40] P. Korshunov, A. George, G. Ozbulak, and S. Marcel, Vulnerability of face age verification to replay attacks, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[41] K. Kotwal, Latent enhancing autoencoder for occluded image classification, *International Conference on Image Processing*, 2024.

[42] K. Kotwal and S. Marcel, Demographic fairness transformer for bias mitigation in face recognition, *IEEE International Joint Conference on Biometrics (IJCB2)*, 2024.

[43] K. Kotwal and S. Marcel, Mitigating demographic bias in face recognition via regularized score calibration, *IEEE/CVF Winter Conference on Applications of Computer Vision Workshops*, 2024.

[44] K. Kotwal, G. Ozbulak, and S. Marcel, Assessing the reliability of biometric authentication on virtual reality devices, *IEEE International Joint Conference on Biometrics*, 2024.

[45] S. Kumar, S. Madikeri, I. Nigmatulina, E. Villatoro-Tello, P. Motlicek, K. Pandia, S. P. Dubagunta, and A. Ganapathiraju, Multitask speech recognition and speaker change detection for unknown number of speakers, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024, pp. 12 592–12 596.

[46] S. Kumar, S. Madikeri, J. Zuluaga-Gomez, I. Thorbecke, E. Villatoro-Tello, S. Burdisso, P. Motlicek, K. P. D. S, and A. Ganapathiraju, Tokenverse: Towards unifying speech and nlp tasks via transducer-based asr, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, ACL, 2024.

[47] C. Li, X. Wu, T. Teng, S. Calinon, and F. Chen, Towards robo-coach: Robot interactive stiffness/position adaptation for human strength and conditioning training, *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[48] C. Li, D. Teney, L. Yang, Q. Wen, X. Xie, and J. Wang, CulturePark: Boosting cross-cultural understanding in large language models, *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

[49] Y. Li, X. Chi, A. Razmjoo, and S. Calinon, Configuration space distance fields for manipulation planning, *Robotics: Science and Systems (RSS)*, 2024.

[50] Y. Li, Y. Zhang, A. Razmjoo, and S. Calinon, Representing robot geometry as distance fields: Applications to whole-body manipulation, *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[51] Y. Linghu, T. de Freitas Pereira, C. Ecabert, S. Marcel, and M. Günther, Score normalization for demographic fairness in face recognition, *IEEE International Joint Conference on Biometrics (IJCB)*, 2024.

[52] J. Liu, C. Li, S. Wang, Z. Dong, Z. Tang, S. Calinon, M. Li, and F. Chen, Learning goal-oriented bimanual dough rolling using dynamic heterogeneous graph based on human demonstration, *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2024.

[53] P. Liu, J. Günster, N. Funk, S. Gröger, D. Chen, H. Bou-Ammar, J. Jankowski, A. Marić, S. Calinon, A. Orsula, M. Olivares-Mendez, H. Zhou, R. Lioutikov, G. Neumann, A. Likmeta, A. Zhalehmehrabi, T. Bonenfant, M. Restelli, D. Tateo, Z. Liu, and J. Peters, A retrospective on the robot air hockey challenge: Benchmarking robust, reliable, and safe learning techniques for real-world robotics, *Neural Information Processing Systems Track on Datasets and Benchmarks (NeurIPS Datasets and Benchmarks)*, 2024.

[54] Z. Liu, W. Sun, Y. Hong, D. Teney, and S. Gould, Bi-directional training for composed image retrieval via text prompt learning, *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024.

[55] T. Löw and S. Calinon, Extending the cooperative dual-task space in conformal geometric algebra, *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[56] T. Löw and S. Calinon, Recursive forward dynamics for serial kinematic chains using conformal geometric algebra, *International Workshop on the Algorithmic Foundations of Robotics (WAFR)*, 2024.

[57] W. Lu, H. Yu, J. Wang, D. Teney, H. Wang, Y. Chen, Q. Yang, X. Xie, and X. Ji, Zoopfl: Exploring black-box foundation models for personalized federated learning, *NeurIPS Workshop on Federated Learning*, 2024.

[58] A. R. Mader, L. B. Meegahapola, and D. Gatica-Perez, Learning about social context from smartphone data: Generalization across countries and daily life moments, *ACM Conference on Human Factors in Computing Systems*, 2024.

[59] F. Marelli and M. Liebling, Optomechanical modulation tomography for ungated compressive cardiac light sheet microscopy, *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2024.

[60] J. Meadows, M. Valentino, D. Teney, and A. Freitas, A symbolic framework for systematic evaluation of mathematical reasoning with transformers, *Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.

[61] M. van der Meer, N. Falk, P. K. Murukannaiah, and E. Liscio, Annotator-centric active learning for subjective nlp tasks, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Association for Computational Linguistics, 2024.

[62] P. Motlicek, E. Dikici, S. Madikeri, P. Rangappa, M. Janosik, G. Backfried, D. Thomas-Aniola, M. Schurz, J. Rohdin, P. Schwarz, M. Kovac, K. Malý, D. Boboš, M. Leibiger, C. Kalogiros, A. Alexopoulos, D. Kudenko, Z. Ahmadi, H. H. Nguyen, A. Krishnan, D. Zhu, D. Klakow, M. Jofre, F. Calderoni, D. Marraud, N. Koutras, N. Nikolau, C. Apostiki, P. Douris, K. Gkountas, E. Sergidou, W. Bosma, J. Hughues, and H. P. Team, Roxsd: The roxanne multimodal and simulated dataset for advancing criminal investigations, *Odyssey 2024: The Speaker and Language Recognition Workshop*, 2024.

[63] M. Naderi, E. Hermann, A. Nanchen, S. Hovsepyan, and M. Magimai-Doss, Towards interfacing large language models with ASR systems using confidence measures and prompting, *Annual Conference of the International Speech Communication Association (Interspeech)*, 2024.

[64] V. Nastase, C. Jiang, G. Samo, and P. Merlo, Exploring syntactic information in sentence embeddings through multilingual subject-verb agreement, *Tenth Italian Conference on Computational Linguistics*, 2024.

[65] V. Nastase and P. Merlo, Are there identifiable structural parts in the sentence embedding whole? *BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP*, 2024.

[66] V. Nastase and P. Merlo, Tracking linguistic information in transformer-based sentence embeddings through targeted sparsification, *Workshop on Representation Learning for NLP*, 2024.

[67] V. Nastase, G. Samo, C. Jiang, and P. Merlo, Exploring Italian sentence embeddings properties through multi-tasking, *Tenth Italian Conference on Computational Linguistics*, 2024.

[68] H. Otroshi Shahreza, C. Ecabert, A. George, A. Unnervik, and S. Marcel, Sdfr: Synthetic data for face recognition competition, *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, 2024.

[69] H. Otroshi Shahreza, A. George, A. Unnervik, P. Rahimi, and S. Marcel, Frcsyn challenge at cvpr: Face recognition challenge in the era of synthetic data, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[70] H. Otroshi Shahreza and S. Marcel, Breaking template protection: Reconstruction of face images from protected facial templates, *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, 2024.

[71] H. Otroshi Shahreza and S. Marcel, Face reconstruction from partially leaked facial embeddings, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[72] H. Otroshi Shahreza and S. Marcel, Unveiling synthetic faces: How synthetic datasets can expose real identities, *NeurIPS Workshop on New Frontiers in Adversarial Machine Learning*, 2024.

[73] H. Otroshi Shahreza, A. Veuthey, and S. Marcel, Face recognition using lensless camera, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[74] Y. L. Pang, C. Oh, and A. Cavallaro, Sparse multi-view hand-object reconstruction for unseen environments, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2024.

[75] A. Pannatier, E. Courdier, and F. Fleuret, Sigma-GPTs: A new approach to autoregressive models, *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2024.

[76] M. Parra, J. Favela, L. Castro, and D. Gatica-Perez, Towards wine tasting activity recognition for a digital sommelier, *ACM Workshop on Exploring Innovative Technology for Commensality and Human-Food Interaction*, 2024.

[77] L. F. Parra-Gallego, T. Purohit, B. Vlasenko, J. R. Orozco-Arroyave, and M. Magimai-Doss, Cross-transfer knowledge between speech and text encoders to evaluate customer satisfaction, *Annual Conference of the International Speech Communication Association (Interspeech)*, ISCA, 2024.

[78] F. Piras, E. De Moura Presa, P. Wellig, and M. Liebling, Parametric point spread function estimation for thermal imaging systems using easy-to-manufacture random pattern targets, *Target and Background Signatures X: Traditional Methods and Artificial Intelligence*, vol. 13199, SPIE, 2024.

[79] A. Prasad, A. Carofilis, G. Vanderreydt, D. Khalil, S. Madikeri, P. Motlicek, and S. Christof, Fine-tuning self-supervised models for language identification using orthonormal constraint, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[80] A. Prasad, S. Madikeri, D. Khalil, P. Motlicek, and S. Christof, Speech and language recognition with low-rank adaptation of pretrained models, *Annual Conference of the International Speech Communication Association (Interspeech)*, 2024.

[81] X. Quan, M. Valentino, L. A. Dennis, and A. Freitas, Enhancing ethical explanations of large language models through iterative symbolic refinement, *Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

[82] X. Quan, M. Valentino, L. A. Dennis, and A. Freitas, Verification and refinement of natural language explanations through llm-symbolic theorem proving, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024.

[83] D. Queiroz Neto, A. Anjos, and L. Berton, Using backbone foundation model for evaluating fairness in chest radiography without demographic data, *IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR)*, 2024.

[84] D. Queiroz Neto, A. Carlos, M. Fatoretto, L. F. Nakayama, A. Anjos, and L. Berton, Does data-efficient generalization exacerbate bias in foundation models? *European Conference on Computer Vision (ECCV)*, 2024.

[85] P. Rahimi, B. Razeghi, and S. Marcel, Synthetic to authentic: Transferring realism to 3D face renderings for boosting face recognition, *European Conference on Computer Vision Workshops*, 2024.

[86] S. Rajapakshe, A. Dastenavar, and E. Senft, Synergizing natural language towards enhanced shared autonomy, *Companion of the ACM/IEEE International Conference on Human-Robot Interaction*, 2024.

[87] N. Ramoly, A. Komaty, V. Krivokuca, L. Younes, A.-M. Awal, and S. Marcel, A novel and responsible dataset for face presentation attack detection on mobile devices, *The IEEE International Joint Conference on Biometrics*, 2024.

[88] P. Rangappa, M. Amand, A. S. Lara, P. Motlicek, M. Antonopoulou, I. Fourfouris, A. Skarlatos, N. Avgerinos, M. Tsangaris, and K. Kostka, Detecting criminal networks via non-content communication data analysis techniques from the tracy project, *EAI International Conference on Digital Forensics & Cyber Crime*, 2024.

[89] B. Razeghi, P. Rahimi, and S. Marcel, Deep variational privacy funnel: General modeling with applications in face recognition, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[90] J. Rozanova, M. Valentino, and A. Freitas, Estimating the causal effects of natural logic features in transformer-based nli models, *Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, 2024.

[91] S. Sajadmanesh and D. Gatica-Perez, Progap: Progressive graph neural networks with differential privacy guarantees, *ACM International Conference on Web Search and Data Mining*, 2024.

[92] D. Sanchez-Cortes, S. Burdisso, E. Villatoro-Tello, and P. Motlicek, Mapping the media landscape: Predicting factual reporting and political bias through web interactions, *International Conference of the CLEF Association*, 2024.

[93] E. Sarkar and M. Magimai-Doss, On the utility of speech and audio foundation models for marmoset call analysis, *International Workshop on Vocal Interactivity in-and-between Humans, Animals and Robots*, 2024.

[94] S. Shakeel and I. Kodrasi, Impact of speech mode in automatic pathological speech detection, *European Signal Processing Conference*, 2024.

[95] S. Shetty, T. Xue, and S. Calinon, Generalized policy iteration using tensor approximation for hybrid control, *International Conference on Learning Representations (ICLR)*, 2024.

[96] S. Tafasca, A. Gupta, V. Bros, and J.-M. Odobez, Toward semantic gaze target detection, *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

[97] S. Tafasca, A. Gupta, and J.-M. Odobez, Sharingan: A transformer architecture for multi-person gaze following, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[98] N. Tarigopula, P. Garg, S. Muralidhar, S. Tornay, D. B. Jayagopi, and M. Magimai-Doss, Content-based objective evaluation of artificially generated sign language videos, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[99] D. Teney, A. M. Nicolicioiu, V. Hartmann, and E. Abbasnejad, Neural redshift: Random networks are not random functions, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[100] D. Teney, J. Wang, and E. Abbasnejad, Selective mixup helps with distribution shifts, but not (only) because of mixup, *International Conference on Machine Learning (ICML)*, 2024.

[101] M. Thayaparan, M. Valentino, and A. Freitas, A differentiable integer linear programming solver for explanation-based natural language inference, *Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, 2024.

[102] I. Thorbecke, J. Zuluaga-Gomez, E. Villatoro-Tello, S. Kumar, P. Rangappa, S. Burdisso, P. Motlicek, K. P. D. S, and A. Ganapathiraju, Fast streaming transducer ASR prototyping via knowledge distillation with Whisper, *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, ACL, 2024.

[103] O. Jimenez-del Toro, C. Aberle, R. Schaer, M. Bach, K. Flouris, E. Konukoglu, B. Stieltjes, M. M. Obmann, A. Anjos, H. Müller, and A. Depeursinge, Comparing stability and discriminatory power of hand-crafted versus deep radiomics: A 3d-printed anthropomorphic phantom study, *European Workshop on Visual Information Processing*, 2024.

[104] A. Unnervik, A. George, C. Ecabert, P. Rahimi, H. Otroshi Shahreza, and S. Marcel, FRCsyn challenge at WACV: Face recognition challenge in the era of synthetic data, *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, 2024.

[105] A. Unnervik, H. Otroshi Shahreza, A. George, and S. Marcel, Model pairing using embedding translation for backdoor attack detection on open-set classification tasks, *Safe Generative AI Workshop at NeurIPS*, 2024.

[106] M. Valentino, D. Carvalho, and A. Freitas, Multi-relational hyperbolic word embeddings from natural language definitions, *Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

[107] L. Vásquez-Rodríguez, B. Audrin, S. Michel, S. Galli, J. Rogenhofer, J. Negro Cusa, and L. van der Plas, A human perspective to AI-based candidate screening, *Hawaii International Conference on System Sciences (HICSS)*, 2024.

[108] L. Vásquez-Rodríguez, B. Audrin, S. Michel, S. Galli, J. Rogenhofer, J. Negro Cusa, and L. van der Plas, Hardware-effective approaches for skill extraction in job offers and resumes, *Workshop on Recommender Systems for Human Resources at the ACM Conference on Recommender Systems*, vol. 3788, 2024.

[109] M. Villamizar, O. Canévet, and J.-M. Odobez, Investigating semantic segmentation models to assist visually impaired people, *European Conference on Computer Vision - Workshops*, 2024.

[110] E. Villatoro-Tello, S. Madikeri, B. Sharma, D. Khalil, S. Kumar, N. Iuliia, P. Motlicek, and A. Ganapathiraju, Probability-aware word-confusion-network-to-text alignment approach for intent classification, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[111] B. Vlasenko, S. Vyas, and M. Magimai-Doss, Comparing data-driven and handcrafted features for dimensional emotion recognition, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2024.

[112] P. Vuillecard, A. Farkhondeh, M. Villamizar, and J.-M. Odobez, CCDb-HG: Novel annotations and gaze-aware representations for head gesture recognition, *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, 2024.

[113] O. Wysocki, M. Wysocka, D. Carvalho, A. T. Bogatu, D. M. Gusicuma, M. Delmas, H. Unsworth, and A. Freitas, An llm-based knowledge synthesis and scientific reasoning framework for biomedical discovery, *Annual Meeting of the Association for Computational Linguistics (ACL), System Demonstrations Track*, 2024.

[114] A. Xompero, M. Bontonou, J.-M. Arbona, E. Benetos, and A. Cavallaro, Explaining models relating objects and privacy, *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2024.

[115] T. Xue, R. Amirreza, and S. Calinon, D-lgp: Dynamic logic-geometric program for reactive task and motion planning, *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[116] T. Xue, A. Razmjoo, S. Shetty, and S. Calinon, Logic-skill programming: An optimization-based approach to sequential skill planning, *Robotics: Science and Systems (RSS)*, 2024.

[117] T. Xue, A. Razmjoo Fard, S. Shetty, and S. Calinon, Robust manipulation primitive learning via domain contraction, *Conference on Robot Learning (CoRL)*, 2024.

[118] Y. Zhang, D. Carvalho, M. Valentino, I. Pratt-Hartmann, and A. Freitas, Improving semantic control in discrete latent spaces with transformer quantized variational autoencoders, *Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

Idiap Research Institute
Rue Marconi 19
1920 Martigny